U.S. DEPARTMENT OF COMMERCE
Patent and Trademark Office

SEARCH REQUEST FORM

Requestor's Name: _____ Serial Number: _____

Date: _____ , Phone: _____ Art Unit: _____

Search Topic:

Please write a detailed statement of search topic. Describe specifically as possible the subject matter to be searched. Define any terms that may have a special meaning. Give examples or relevent citations, authors, keywords, etc., if known. For sequences, please attach a copy of the sequence. You may include a copy of the broadest and/or most relevent claim(s).

STAFF USE ONLY

Date completed: 07-15-02
Searcher: Beverly E 4994
Terminal time: 20
Elapsed time: _____
CPU time: _____
Total time: 25
Number of Searches: _____
Number of Databases: 1

Search Site

_____ STIC
_____ CM-1
_____ Pre-S

Type of Search

_____ N.A. Sequence
_____ A.A. Sequence
_____ Structure
_____ Bibliographic

Vendors

_____ IG
_____ STN
_____ Dialog
_____ APS
_____ Geninfo
_____ SDC
_____ DARC/Questel
_____ Other CGN

PTO-1590 (9-90)

OM protein - protein search, using sw model

Run on: July 15, 2002, 12:16:23 ; Search time 58.93 Seconds
(without alignments)
1053.627 Million cell updates/sec

Title: US-09-298-924-6
Perfect score: 2978
Sequence: 1 MTFAYKIDGNEVIFTLWAPY.....PQHIEEGKYEFDKGFALYKL 559

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802:*
1: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2978	100.0	559	17	AAR90619	Sulfolobus solfata
2	1748.5	58.7	556	17	AAR92755	Trehalose-releasin
3	1748.5	58.7	556	17	AAR90620	Sulfolobus acidoca
4	1089.5	36.6	595	22	AAG92072	C glutamicum prote
5	1089.5	36.6	610	22	AAB80206	Corynebacterium gl
6	1060.5	35.6	598	16	AAR77471	Trehalose releasin
7	1058	35.5	597	16	AAR80290	Trehalose releasin
8	1025	34.4	589	16	AAR80289	Trehalose releasin
9	1025	34.4	596	16	AAR77470	Trehalose releasin
10	1024.5	34.4	575	21	AAY85157	Trehalose-releasin
11	1024.5	34.4	575	21	AAY85165	Trehalose-releasin
12	399	13.4	652	16	AAR80037	Bacillus stearothe
13	399	13.4	652	17	AAR96109	Starch-branching-e
14	385.5	12.9	648	10	AAP94635	B. thuringiensis a
15	374.5	12.6	639	12	AAR11271	B.stearothermophil
16	371.5	12.5	718	22	AAU60903	Propionibacterium
17	361	12.1	630	22	AAB69074	Aquifex aeolicus V
18	348	11.7	670	22	AAU60913	Propionibacterium
19	348	11.7	766	20	AAY00870	S. tuberosum isoam
20	346.5	11.6	658	13	AAR23787	Heat-resistant pul
21	345.5	11.6	666	20	AAY34991	C. pneumoniae prot
22	342.5	11.5	793	20	AAY00869	S. tuberosum isoam
23	340	11.4	731	22	AAG91100	C glutamicum prote
24	340	11.4	731	22	AAB79423	Corynebacterium gl
25	339	11.4	764	21	AAY50819	Wheat isoamylase p
26	334	11.2	606	18	AAW36602	S. tuberosum debra
27	333	11.2	1250	21	AAY91279	Group B Streptococ
28	331.5	11.1	618	20	AAY37184	Protein involved i
29	331.5	11.1	798	20	AAW73552	Pullulanase protei
30	331.5	11.1	1938	20	AAW73553	Full length Pullul
31	329.5	11.1	818	20	AAY17523	Zea mays SU1 starc
32	329.5	11.1	893	17	AAW09257	Bacillus alkaline
33	329.5	11.1	1938	17	AAW09255	Bacillus alkaline
34	315	10.6	772	19	AAW49871	Thermotoga maritim
35	312.5	10.5	621	21	AAB19285	A polypeptide with
36	312	10.5	772	18	AAW34567	Thermotoga maritim
37	309.5	10.4	722	20	AAY35095	C. pneumoniae prot
38	303.5	10.2	829	22	AAE05691	Bacillus acidopull
39	303.5	10.2	862	21	AAY78513	Pullulanase (pulB)
40	303	10.2	931	20	AAY27357	Group B Streptococ
41	300	10.1	498	10	AAP96105	Sequence of amy C
42	299	10.0	928	22	AAE05719	Bacillus deramific
43	296	9.9	762	21	AAW90977	N. denitrificans a
44	295	9.9	555	14	AAR43340	Alpha-glucosidase
45	294.5	9.9	785	19	AAW70886	Protein encoded by

ALIGNMENTS

RESULT 1
AAR90619
ID AAR90619 standard; Protein; 559 AA.
XX
AC AAR90619;
XX
DT 29-JUN-1996 (first entry)
XX
DE Sulfolobus solfataricus amylase for alpha, alpha-trehalose prodn.
XX
KW transferase; amylase; Sulfolobus; production; alpha, alpha-trehalose;
KW malto-oligosaccharide; hydrolysis.
XX
OS Sulfolobus solfataricus.
XX
PN WO9534642-A.
XX
PD 21-DEC-1995.
XX
PF 14-JUN-1995; 95WO-JP01189.
XX
PR 21-APR-1995; 95JP-0120673.
PR 15-JUN-1994; 94JP-0133354.
PR 18-AUG-1994; 94JP-0194223.
PR 31-OCT-1994; 94JP-0290394.
PR 21-NOV-1994; 94JP-0286917.
PR 21-NOV-1994; 94JP-0311185.
XX
PA (KIRI) KIRIN BEER KK.
XX
DR WPI; 1996-049671/05.
DR N-PSDB; AAT12325.
XX
PT Sulfolobus spp. derived transferase and amylase - for production of

```
PT   alpha, alpha-trehalose from malto-oligosaccharide(s)
XX
PS   Claim 108; Page 235-240; 357pp; Japanese.
XX
CC   The amylase is derived from Sulfolobus solfataricus.  The amylase acts
CC   on a saccharide having at least three sugar units, which are pref.
CC   glucose units at the reducing end (the linkage between the first and
CC   second glucose units is alpha-1, alpha-1, while the linkage between the
CC   second and third glucose units is alpha-1,4), to hydrolyse alpha-1,4
CC   linkages within the sugar chain, yielding alpha, alpha-trehalose and also
CC   mono- and disaccharide hydrolysis products. The amylase has a mol. wt. of
CC   61 to 64 kDa.  It is characterised by working at pH 4.5-5.5 and at 60-85
CC   deg.C.  It has an isoelectric point of 4.3-5.4 and retains at least
CC   100 percent activity after 6 hrs. at 80 deg.C. It is completely inhibited
CC   by 5 mM copper sulphate. Use of a transferase and the amylase in
CC   succession on suitable substrates such a malto-oligosaccharides, is
CC   useful for the production of alpha, alpha-trehalose.
XX
SQ   Sequence   559 AA;


  Query Match             100.0%;  Score 2978;  DB 17;  Length 559;
  Best Local Similarity   100.0%;  Pred. No. 3.6e-236;
  Matches  559;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MTFAYKIDGNEVIFTLWAPYQKSVKLKVLEKGLYEMERDEKGYFTITLNNVKVRDRYKYV 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 mtfaykidgneviftlwapyqksvklkvlekglyemerdekgyftitlnnvkvrdrykyv 60

Qy      61 LDDASEIPDPASRYQPEGVHGPSQIIQESKEFNNETFLKKEDLIIYEIHVGTFTPEGTFE 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 lddaseipdpasryqpegvhgpsqiiqeskefnnetflkkedliiyeihvgtftpegtfe 120

Qy     121 GVIRKLDYLKDLGITAIEIMPIAQFPGKRDWGYDGVYLYAVQNSYGGPEGFRKLVDEAHK 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 gvirkldylkdlgitaieimpiaqfpgkrdwgydgvylyavqnsyggpegfrklvdeahk 180

Qy     181 KGLGVILDVVYNHVGPEGNYMVKLGPYFSQKYKTPWGLTFNFDDAESDEVRKFILENVEY 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 kglgvildvvynhvgpegnymvklgpyfsqkyktpwgltfnfddaesdevrkfilenvey 240

Qy     241 WIKEYNVDGFRLDAVHAIIDTSPKHILEEIADVVHKYNRIVIAESDLNDPRVVNPKEKCG 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 wikeynvdgfrldavhaiidtspkhileeiadvvhkynriviaesdlndprvvnpkekcg 300

Qy     301 YNIDAQWVDDFHHSIHAYLTGERQGYYTDFGNLDDIVKSYKDVFVYDGKYSNFRRKTHGE 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 ynidaqwvddfhhsihayltgerqgyytdfgnlddivksykdvfvydgkysnfrrkthge 360

Qy     361 PVGELDGCNFVVYIQNHDQVGNRGKGERIIKLVDRESYKIAAALYLLSPYIPMIFMGEEY 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 pvgeldgcnfvvyiqnhdqvgnrgkgeriiklvdresykiaaalyllspyipmifmgeey 420

Qy     421 GEENPFYFFSDFSDSKLIQGVREGRKKENGQDTDPQDESTFNASKLSWKIDEEIFSFYKI 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 geenpfyffsdfsdskliqgvregrkkengqdtdpqdestfnasklswkideeifsfyki 480

Qy     481 LIKMRKELSIACDRRVNVVNGENWLIIKGREYFSLYVFSKSSIEVKYSGTLLLSSNNSFP 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 likmrkelsiacdrrvnvvngenwliikgreyfslyvfskssievkysgtlllssnnsfp 540

Qy     541 QHIEEGKYEFDKGFALYKL 559
           |||||||||||||||||||
Db     541 qhieegkyefdkgfalykl 559


RESULT   2
AAR92755
ID   AAR92755 standard; Protein; 556 AA.
XX
AC   AAR92755;
XX
DT   03-AUG-1996  (first entry)
XX
DE   Trehalose-releasing thermostable enzyme.
XX
KW   Thermostable enzyme; trehalose; sweetener.
XX
OS   Sulfolobus acidocaldarius strain ATCC 33909.
XX
PN   AU9527131-A.
XX
PD   01-FEB-1996.
XX
PF   21-JUL-1995;   95AU-0027131.
XX
PR   04-JUL-1995;   95JP-0189760.
PR   21-JUL-1994;   94JP-0190180.
PR   11-APR-1995;   95JP-0109128.
XX
PA   (HAYB ) HAYASHIBARA SEIBUTSU KAGAKU.
XX
PI   Kubota M,  Mitsuzumi H,  Sugimoto T;
XX
DR   WPI; 1996-106284/12.
DR   N-PSDB; AAT16899.
XX
PT   Recombinant thermostable enzyme from Sulfolobus acidocaldarius,  -
PT   releases trehalose from non-reducing saccharide at temps. exceeding
PT   55 degrees Centigrade
XX
PS   Claim 2; Page 53-54; 74pp; English.
XX
CC   A thermostable enzyme (AAR92755) of Sulfolobus acidocaldarius ATCC
CC   33909 releases trehalose from non-reducing saccharides having a
CC   trehalose structure as an end unit and a degree of polymerisation of
CC   at least 3.  It has a mol.wt. of 54-64 kDa (SDS-PAGE), a pI of
CC   5.6-6.6 and is substantially not inactivated when incubated in aq.
CC   solution (pH 7.0) at 85 deg for 60 min.  Recombinant enzyme is
CC   obtd. by expression of an isolated DNA fragment (AAT16899) in
CC   host cells, pref. Escherichia coli, using e.g. vector pBluescript
CC   II SK(+).  The trehelose is useful as a sweetener.
XX
SQ   Sequence   556 AA;


  Query Match             58.7%;  Score 1748.5;  DB 17;  Length 556;
  Best Local Similarity   58.9%;  Pred. No. 4.1e-135;
  Matches  329;  Conservative   87;  Mismatches  136;  Indels    7;  Gaps    5;

Qy       2 TFAYKIDGNEVIFTLWAPYQKSVKLKVLEKGLYEMERDEKGYFTITLNNVKVRDRYKYVL 61
           :|   |: |: || |||||  |||| | | |  ||:::::|:| : :::::    | |::
Db       3 sfggniekngkifklwapyvnsvklk-lskklipmeknegffeveiddieenltysyii 61

Qy      62 DDASEIPDPASRYQPEGVHGPSQIIQESKEFNNETFLKKEDLIIYEIHVGTFTPEGTFEG 121
           :|  ||||||||||| |||  ||:|:   :  :   :| |||||||:|||||: || |:|
Db      62 edkreipdpasryqplgvhdksqlirtdyqildlgkvkiedliiyelhvgtfsqegnfkg 121

Qy     122 VIRKLDYLKDLGITAIEIMPIAQFPGKRDWGYDGVYLYAVQNSYGGPEGFRKLVDEAHKK 181
           || |||||||||| ||:||:||||| |||||||:||||||:||||    |||:||||:
Db     122 viekldylkdlgitgielmpvaqfpgnrdwgydgvflyavqntyggpewelaklvneahkr 181

Qy     182 GLGVILDVVYNHVGPEGNYMVKLGPYFSQKYKTPWGLTFNFDDAESDEVRKFILENVEYW 241
           |: ||||||||:|||||:: |||||| :|||||||||||||   |:|||||||||||
Db     182 giavildvvynhigpegnyllglgpyfsdryktpwgltfnfddrgcdqvrkfilenveyw 241

Qy     242 IKEYNVDGFRLDAVHAIIDTSPKHILEEIADVVHKYNRIVIAESDLNDPRVVNPKEKCGY 301
           | : :|| |||||||| | ||||||:|||:  |:  : ||||||||||::|  |: |||
Db     242 fktfkidglrldavhaifdnspkhilqeiaekahqlgkfviaesdlndpkiv--kddcgy 299

Qy     302 NIDAQWVDDFHHSIHAYLTGERQGYYTDFGNLDDIVKSYKDVFVYDGKYSNFRRKTHGEP 361
           ||||||||||::||::| |:  || ||| ::|| |::||||||||||| :| :||| |
Db     300 kidaqwvddfhhavhafitkekdyyyqdfgriediektfkdvfvydgkysryrgrthgap 359

Qy     362 VGELDGCNFVVYIQNHDQVGNRGKGERIIKLVDRESYKIAAALYLLSPYIPMIFMGEEYG 421
           ||:|    |||:|||||||||| |||:  | |: :| :|| ||:|||||:|||||||
```

```
Db    360 vgdlpprkfvvfiqnhdqvgnrgngerlsiltdkttylmaatlyilspyiplifmgeeyy 419

Qy    422 EENPFYFFSDFSDSKLIQGVREGRKKENGQDTDPQDESTFNASKLSWKIDEEIFSFYKIL 481
          | |||:|||||||  ||:|||||| ||| |  ||| |  |  |||||||||:  :|| |
Db    420 etnpfffsdfsdpvlikgvregrlkennqmidpqseeaflksklswkideevldyykql 479

Qy    482 IKMRKELSIACDRRVNVVNGENWLIIKGREYFSLYVFSKSSIEVKYSGTLLLSSNNSFPQ 541
          | :||  :  | |   |    | : :   :   :  |    |  | :| ||:    ||:
Db    480 inirkryn-nckrvkevrregncitlimekigiiasfddivinskitgnlligi--gfpk 536

Qy    542 HIEEGK-YEFDKGFALYKL 559
           ::: :  : ::|  :|:|
Db    537 klkkdelikvnrgvgvyql 555


RESULT   3
AAR90620
ID   AAR90620 standard; Protein; 556 AA.
XX
AC   AAR90620;
XX
DT   29-JUN-1996  (first entry)
XX
DE   Sulfolobus acidocaldarius amylase for alpha, alpha-trehalose prodn.
XX
KW   transferase; amylase; Sulfolobus; production; alpha, alpha-trehalose;
KW   malto-oligosaccharide; hydrolysis.
XX
OS   Sulfolobus acidocaldarius.
XX
PN   WO9534642-A.
XX
PD   21-DEC-1995.
XX
PF   14-JUN-1995;   95WO-JP01189.
XX
PR   21-APR-1995;   95JP-0120673.
PR   15-JUN-1994;   94JP-0133354.
PR   18-AUG-1994;   94JP-0194223.
PR   31-OCT-1994;   94JP-0290394.
PR   21-NOV-1994;   94JP-0286917.
PR   21-NOV-1994;   94JP-0311185.
XX
PA   (KIRI ) KIRIN BEER KK.
XX
DR   WPI; 1996-049671/05.
DR   N-PSDB; AAT12326.
XX
PT   Sulfolobus spp. derived transferase and amylase - for production of
PT   alpha, alpha-trehalose from malto-oligosaccharide(s)
XX
PS   Claim 112; Page 244-250; 357pp; Japanese.
XX
CC   The amylase is derived from Sulfolobus acidocaldarius.  The amylase acts
CC   on a saccharide having at least three sugar units, which are pref.
CC   glucose units at the reducing end (the linkage between the first and
CC   second glucose units is alpha-1, alpha-1, while the linkage between the
CC   second and third glucose units is alpha-1,4), to hydrolyse alpha-1,4
CC   linkages within the sugar chain, yielding alpha, alpha-trehalose and also
CC   mono- and disaccharide hydrolysis products. The amylase has a mol. wt. of
CC   61 to 64 kDa.  It is characterised by working at pH 4.5-5.5 and at 60-85
CC   deg.C.  It has an isoelectric point of 4.3-5.4 and retains at least
CC   100 percent activity after 6 hrs. at 80 deg.C. It is completely inhibited
CC   by 5 mM copper sulphate. Use of a transferase and the amylase in
CC   succession on suitable substrates such a malto-oligosaccharides, is
CC   useful for the production of alpha, alpha-trehalose.
XX
SQ   Sequence   556 AA;


  Query Match             58.7%;  Score 1748.5;  DB 17;  Length 556;
  Best Local Similarity   58.9%;  Pred. No. 4.1e-135;
  Matches  329;  Conservative   87;  Mismatches  136;  Indels    7;  Gaps    5;

Qy      2 TFAYKIDGNEVIFTLWAPYQKSVKLKVLEKGLYEMERDEKGYFTITLNNVKVRDRYKYVL 61
            :|   |: |: || |||||  |||||  | | |  ||::::|:| : :::::    | |::
Db      3 sfggnieknkgifklwapyvnsvklk-lskklipmekndegffeveiddieenltysyii 61

Qy     62 DDASEIPDPASRYQPEGVHGPSQIIQESKEFNNETFLKKEDLIIYEIHVGTFTPEGTFEG 121
            :|  ||||||||||| |||  ||:|:   :  :   :| |||||||:|||||: || |:|
Db     62 edkreipdpasryqplgvhdksqlirtdyqildlgkvkiedliiyelhvgtfsqegnfkg 121

Qy    122 VIRKLDYLKDLGITAIEIMPIAQFPGKRDWGYDGVYLYAVQNSYGGPEGFRKLVDEAHKK 181
            || |||||||||| ||:||:||||| ||||||||:||||||:||||    |||:||||:
Db    122 viekldylkdlgitgielmpvaqfpgnrdwgydgvflyavqntyggpwelaklvneahkr 181

Qy    182 GLGVILDVVYNHVGPEGNYMVKLGPYFSQKYKTPWGLTFNFDDAESDEVRKFILENVEYW 241
            |: |||||||||:||||||:: |||||| :|||||||+||||||   |:||||||||||||
Db    182 giavildvvynhigpegnyllglgpyfsdryktpwgltfnfddrgcdqvrkfilenveyw 241

Qy    242 IKEYNVDGFRLDAVHAIIDTSPKHILEEIADVVHKYNRIVIAESDLNDPRVVNPKEKCGY 301
            | : :|| |||||||| | ||||||:|||:  |:  : ||||||||||::|  |: |||
Db    242 fktfkidglrldavhaifdnspkhilqeiaekahqlgkfviaesdlndpkiv--kddcgy 299

Qy    302 NIDAQWVDDFHHSIHAYLTGERQGYYTDFGNLDDIVKSYKDVFVYDGKYSNFRRKTHGEP 361
            ||||||||||||::||::| |:  || ||| ::|| |::||||||||||| :| :||| |
Db    300 kidaqwvddfhhavhafitkekdyyyqdfgriediektfkdvfvydgkysryrgrthgap 359

Qy    362 VGELDGCNFVVYIQNHDQVGNRGKGERIIKLVDRESYKIAAALYLLSPYIPMIFMGEEYG 421
            ||:|    |||:|||||||||| |||:  | |: :| :|| ||:|||||:|||||||
Db    360 vgdlpprkfvvfiqnhdqvgnrgngerlsiltdkttylmaatlyilspyiplifmgeeyy 419

Qy    422 EENPFYFFSDFSDSKLIQGVREGRKKENGQDTDPQDESTFNASKLSWKIDEEIFSFYKIL 481
            | |||:|||||||  ||:|||||| ||| |  ||| |  |  |||||||||:  :|| |
Db    420 etnpfffsdfsdpvlikgvregrlkennqmidpqseeaflksklswkideevldyykql 479

Qy    482 IKMRKELSIACDRRVNVVNGENWLIIKGREYFSLYVFSKSSIEVKYSGTLLLSSNNSFPQ 541
            | :||  :  | |   |    | : :   :   :  |    |  | :| ||:    ||:
Db    480 inirkryn-nckrvkevrregncitlimekigiiasfddivinskitgnlligi--gfpk 536

Qy    542 HIEEGK-YEFDKGFALYKL 559
           ::: :  : ::|  :|:|
Db    537 klkkdelikvnrgvgvyql 555


RESULT   4
AAG92072
ID   AAG92072 standard; Protein; 595 AA.
XX
AC   AAG92072;
XX
DT   26-SEP-2001  (first entry)
XX
DE   C glutamicum protein fragment SEQ ID NO: 5826.
XX
KW   Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
KW   organic acid synthesis.
XX
OS   Corynebacterium glutamicum.
XX
PN   EP1108790-A2.
XX
PD   20-JUN-2001.
XX
PF   18-DEC-2000; 2000EP-0127688.
XX
PR   16-DEC-1999;   99JP-0377484.
PR   07-APR-2000; 2000JP-0159162.
PR   03-AUG-2000; 2000JP-0280988.
XX
PA   (KYOW ) KYOWA HAKKO KOGYO KK.
XX
PI   Nakagawa S,  Mizoguchi H,  Ando S,  Hayashi M,  Ochiai K,  Yokoi H;
PI   Tateishi N,  Senoh A,  Ikeda M,  Ozaki A;
XX
DR   WPI; 2001-376931/40.
```

```
DR   N-PSDB; AAH67291.
XX
PT   Novel polynucleotides derived from Coryneform bacteria, for identifying
PT   mutation point of a gene, measuring expression of a gene, analysing
PT   expression profile or pattern of a gene and identifying homologous gene
PT   -
XX
PS   Claim 17; SEQ ID NO: 5826; 246pp + Sequence Listing; English.
XX
CC   The present invention provides a number of nucleotide and protein
CC   sequences from the Coryneform bacterium Corynebacterium glutamicum. These
CC   are useful for identifying the mutation point of a gene derived from a
CC   mutant of coryneform bacterium, measuring expression amount and
CC   analysing the expression profile or expression pattern of a gene derived
CC   from Coryneform bacterium, and identifying a homologue of a gene derived
CC   from coryneform bacterium. Coryneform bacteria are useful for producing
CC   amino acids, nucleic acids, vitamins, saccharides and organic acids,
CC   particularly L-lysine. The present sequence is a protein described
CC   in the exemplification of the invention.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from the
CC   European Patent Office.
XX
SQ   Sequence   595 AA;


  Query Match           36.6%;  Score 1089.5;  DB 22;  Length 595;
  Best Local Similarity  41.5%;  Pred. No. 6.5e-81;
  Matches  227;  Conservative   93;  Mismatches  198;  Indels   29;  Gaps   11;

Qy      14 FTLWAPYQKSVKLKVLEKGLYEMERDEKGYFTITLNNVKVRDRYKYVLDDASE----IPD 69
           |::|||    | | :|      | : |  ::   :   |  ||| : | | |     :||
Db      28 fsvwaplphdvhl-ilngetlpmhktegswwraei-apkagdrygfslfdgsswsktlpd 85

Qy      70 PASRYQPEGVHGPSQIIQESKEFNNETFLKK--EDLIIYEIHVGTFTPEGTFEGVIRKLD 127
           | |  ||:|||| |::  :|  : :: :  :     ::||:|||||: :||||||: ||
Db      86 prstsqpdgvhglsevsddsylwgdqqwtgrilpgsvlyelhvgtfsedgtfegvvdklp 145

Qy     128 YLKDLGITAIEIMPIAQFPGKRDWGYDGVYLYAVQNSYGGPEGFRKLVDEAHKKGLGVIL 187
           ||:|||:||||::|:  | | |:|||||  :||   |||| | :||:| :|: |: | |
Db     146 ylrdlgvtaiellpvqpfggnrnwgydgvlwhavhagyggpaglkklidashqagiavyl 205

Qy     188 DVVYNHVGPEGNYMVKLGPYFSQKYKTPWGLTFNFDDAESDEVRKFILENVEYWIKEYNV 247
           |||||| ||:|||  : ||| |    | ||   | :  :|||||  :||:    | ::::|
Db     206 dvvynhfgpdgnyngqfgpytsgg-stgwgdvvninghdsdevrnyildaarqwfedfhv 264

Qy     248 DGFRLDAVHAIIDTSPKHILEEIADVVHKYN------RIVIAESDLNDPRVVNPKEKCGY 301
           || ||||||:: |     :| ::  |    :      | :||||:||||: |  :|  |:
Db     265 dglrldavhsldrgaysllaqltmvaedvsaqtgiprsliaeselndpkfvtsreaggf 324

Qy     302 NIDAQWVDDFHHSIHAYLTGERQGYYTDFGNLDDIVKSYKDVFVYDGKYSNFRRKTHGEP 361
            :|||||| ||::|| ::||| |||:|||::| : |: ::|| : | || :| : || |
Db     325 gldaqwvddihhalhalvsgerngyysdfgsvdtlaktlrevfehtgnystyrgrnhgrp 384

Qy     362 VGE--LDGCNFVVYIQNHDQVGNRGKGERIIKLVDRESYKIAAALYLLSPYIPMIFMGEE 419
           |       || |   ||| ||| |:|    :  |   : ||:   ||| ||:|||||
Db     385 vhpditpasrfvtytthdqtgnraigdrpsttltpeqqvlkaaiiyysspytpmlfmgee 444

Qy     420 YGEENPFYFFSDFSDSKLIQGVREGRKKE------NGQD-TDPQDESTFNASKLSWKIDE 472
           :|   || ||   :| :| :   ||||:|      |  |   |: |||| :||| |:
Db     445 fgattpfaffcshtdpelnrltsegrkrefarlgwnaddipspelestftsskldwefta 504

Qy     473 E---IFSFYKILIKMRKELSIACDR--RVNVVNGENWLIIKGREYFSLYVFSKSSIEVKY 527
           |   |   || |: :|  |  :      : | :|||||  :       |  ||  :| |
Db     505 eqrrindaykqllhlrhtlgfsqpnlltlevehgenwlsmangrgrilanfsddtitvpl 564

Qy     528 SGTLLLS 534
             | |: |
Db     565 ggeliys 571


RESULT   5
AAB80206
```

```
ID   AAB80206 standard; Protein; 610 AA.
XX
AC   AAB80206;
XX
DT   30-APR-2001  (first entry)
XX
DE   Corynebacterium glutamicum MP protein sequence SEQ ID NO:1146.
XX
KW   Corynebacterium glutamicum; metabolic pathway protein; MP protein;
KW   fine chemical production; microorganism; organic acid; nucleoside;
KW   nonproteinogenic amino acid; purine base; pyrimidine base; nucleotide;
KW   lipid; saturated fatty acid; unsaturated fatty acid; diol; vitamin;
KW   carbohydrate; aromatic compound; cofactor; polyketide; enzyme.
XX
OS   Corynebacterium glutamicum.
XX
PN   WO200100843-A2.
XX
PD   04-JAN-2001.
XX
PF   23-JUN-2000; 2000WO-IB00923.
XX
PR   25-JUN-1999;   99US-0141031.
PR   01-JUL-1999;   99DE-1030476.
PR   02-JUL-1999;   99US-0142101.
PR   08-JUL-1999;   99DE-1031415.
PR   08-JUL-1999;   99DE-1031418.
PR   08-JUL-1999;   99DE-1031419.
PR   08-JUL-1999;   99DE-1031420.
PR   08-JUL-1999;   99DE-1031424.
PR   08-JUL-1999;   99DE-1031428.
PR   08-JUL-1999;   99DE-1031434.
PR   08-JUL-1999;   99DE-1031435.
PR   08-JUL-1999;   99DE-1031443.
PR   08-JUL-1999;   99DE-1031453.
PR   08-JUL-1999;   99DE-1031457.
PR   08-JUL-1999;   99DE-1031465.
PR   08-JUL-1999;   99DE-1031478.
PR   08-JUL-1999;   99DE-1031510.
PR   08-JUL-1999;   99DE-1031541.
PR   08-JUL-1999;   99DE-1031573.
PR   08-JUL-1999;   99DE-1031592.
PR   08-JUL-1999;   99DE-1031632.
PR   08-JUL-1999;   99DE-1031634.
PR   08-JUL-1999;   99DE-1031636.
PR   09-JUL-1999;   99DE-1032125.
PR   09-JUL-1999;   99DE-1032126.
PR   09-JUL-1999;   99DE-1032130.
PR   09-JUL-1999;   99DE-1032186.
PR   09-JUL-1999;   99DE-1032206.
PR   09-JUL-1999;   99DE-1032227.
PR   09-JUL-1999;   99DE-1032228.
PR   09-JUL-1999;   99DE-1032229.
PR   09-JUL-1999;   99DE-1032230.
PR   14-JUL-1999;   99DE-1032922.
PR   14-JUL-1999;   99DE-1032926.
PR   14-JUL-1999;   99DE-1032928.
PR   14-JUL-1999;   99DE-1033004.
PR   14-JUL-1999;   99DE-1033005.
PR   14-JUL-1999;   99DE-1033006.
PR   12-AUG-1999;   99US-0148613.
PR   27-AUG-1999;   99DE-1040764.
PR   27-AUG-1999;   99DE-1040765.
PR   27-AUG-1999;   99DE-1040766.
PR   27-AUG-1999;   99DE-1040832.
PR   31-AUG-1999;   99DE-1041378.
PR   31-AUG-1999;   99DE-1041379.
PR   31-AUG-1999;   99DE-1041380.
PR   31-AUG-1999;   99DE-1041394.
PR   31-AUG-1999;   99DE-1041396.
PR   03-SEP-1999;   99DE-1042076.
PR   03-SEP-1999;   99DE-1042077.
PR   03-SEP-1999;   99DE-1042079.
```

```
PR   03-SEP-1999;   99DE-1042086.
PR   03-SEP-1999;   99DE-1042087.
PR   03-SEP-1999;   99DE-1042088.
PR   03-SEP-1999;   99DE-1042095.
PR   03-SEP-1999;   99DE-1042124.
PR   03-SEP-1999;   99DE-1042129.
PR   09-MAR-2000; 2000US-0187970.
XX
PA   (BADI ) BASF AG.
XX
PI   Pompejus M,  Kroeger B,  Schroeder H,  Zelder O,  Haberhauer G;
XX
DR   WPI; 2001-137957/14.
DR   N-PSDB; AAF72325.
XX
PT   Nucleic acids from Corynebacterium glutamicum encoding metabolic
PT   pathway proteins, useful for producing fine chemicals in
PT   microorganisms, including organic acids, nonproteinogenic amino acids,
PT   and purine and pyrimidine bases.-
XX
PS   Claim 20; Page 1722-1725; 1737pp; English.
XX
CC   AAF71753 to AAF72330 encode the Corynebacterium glutamicum metabolic
CC   pathway (MP) proteins given in AAB79634 to AAB80211. The C. glutamicum
CC   MP nucleic acids are useful for the production of fine chemicals
CC   in microorganisms, including organic acids, nonproteinogenic amino
CC   acids, purine and pyrimidine bases, nucleosides, nucleotides, lipids,
CC   saturated and unsaturated fatty acids, diols, carbohydrates, aromatic
CC   compounds, vitamins, cofactors, polyketides and enzymes.
XX
SQ   Sequence   610 AA;
```

```
  Query Match             36.6%;  Score 1089.5;  DB 22;  Length 610;
  Best Local Similarity   41.5%;  Pred. No. 6.7e-81;
  Matches  227;  Conservative   93;  Mismatches  198;  Indels   29;  Gaps   11;

Qy      14 FTLWAPYQKSVKLKVLEKGLYEMERDEKGYFTITLNNVKVRDRYKYVLDDASE----IPD 69
           |::|||    | | :|      | : |  ::   :   |  ||| : | | |     :||
Db      43 fsvwaplphdvhl-ilngetlpmhktegswwraei-apkagdrygfslfdgsswsktlpd 100

Qy      70 PASRYQPEGVHGPSQIIQESKEFNNETFLKK--EDLIIYEIHVGTFTPEGTFEGVIRKLD 127
           | |  ||:|||| |::  :|  : :: :  :     ::||:|||||: :||||||: ||
Db     101 prstsqpdgvhglsevsddsylwgdqqwtgrilpgsvlyelhvgtfsedgtfegvvdklp 160

Qy     128 YLKDLGITAIEIMPIAQFPGKRDWGYDGVYLYAVQNSYGGPEGFRKLVDEAHKKGLGVIL 187
           ||:|||:||||::|:  | | |:|||||  :||   |||| | :||:| :|: |: | |
Db     161 ylrdlgvtaiellpvqpfggnrnwgydgvlwhavhagyggpaglkklidashqagiavyl 220

Qy     188 DVVYNHVGPEGNYMVKLGPYFSQKYKTPWGLTFNFDDAESDEVRKFILENVEYWIKEYNV 247
           |||||| ||:|||  : ||| |    | ||   | :  :|||||  :|||:    | ::::|
Db     221 dvvynhfgpdgnyngqfgpytsgg-stgwgdvvninghdsdevrnyildaarqwfedfhv 279

Qy     248 DGFRLDAVHAIIDTSPKHILEEIADVVHKYN------RIVIAESDLNDPRVVNPKEKCGY 301
           || |||||:: |     :| ::  |    :      | :||||:||||: |  :|  |:
Db     280 dglrldavhsldderqaysllaqltmvaedvsaqtgiprsliaeselndpkfvtsreaggf 339

Qy     302 NIDAQWVDDFHHSIHAYLTGERQGYYTDFGNLDDIVKSYKDVFVYDGKYSNFRRKTHGEP 361
            :|||||| ||::|| ::||| |||:|||::| : |: ::|| : | || :| : || |
Db     340 gldaqwvddihhalhalvsgerngyysdfgsvdtlaktlrevfehtgnystyrgrnhgrp 399

Qy     362 VGE--LDGCNFVVYIQNHDQVGNRGKGERIIKLVDRESYKIAAALYLLSPYIPMIFMGEE 419
           |         || |   ||| |||  |:|    :  |   : ||:   ||| ||:|||||
Db     400 vhpditpasrfvtytttthdqtgnraigdrpsttltpeqqvlkaaiiysspytpmlfmgee 459

Qy     420 YGEENPFYFFSDFSDSKLIQGVREGRKKE------NGQD-TDPQDESTFNASKLSWKIDE 472
           :|   || ||   :| :| :   ||||:|      |  |   |: |||| :||| |:
Db     460 fgattpfaffcshtdpelnrltsegrkrefarlgwnaddipspelestftsskldwefta 519

Qy     473 E---IFSFYKILIKMRKELSIACDR--RVNVVNGENWLIIKGREYFSLYVFSKSSIEVKY 527
           |   |   || |: :|  |  :      : | :|||||  :       |  ||  :| |
Db     520 eqrrindaykqllhlrhtlgfsqpnlltlevehgenwlsmangrgrilanfsddtitvpl 579
```

```
Qy     528 SGTLLLS 534
           | |: |
Db     580 ggeliys 586
```

```
RESULT   6
AAR77471
ID   AAR77471 standard; Protein; 598 AA.
XX
AC   AAR77471;
XX
DT   25-JAN-1996  (first entry)
XX
DE   Trehalose releasing enzyme.
XX
KW   Enzyme; trehalose; alpha-glucosyltrehalose; alpha-maltosyltrehalose;
KW   alpha-maltotriosyltrehalose; alpha-maltotetrasyltrehalose;
KW   maltopentaosyltrehalose; sweetener; taste-improving agent;
KW   stabiliser; filler; excipient; adjuvant; foodstuffs; cosmetics;
KW   pharmaceuticals.
XX
OS   Arthrobacter sp. Q36.
XX
PN   EP671470-A2.
XX
PD   13-SEP-1995.
XX
PF   07-MAR-1995;   95EP-0301474.
XX
PR   07-MAR-1994;   94JP-0059840.
PR   07-MAR-1994;   94JP-0059834.
XX
PA   (HAYB ) HAYASHIBARA SEIBUTSU KAGAKU.
XX
PI   Hattori K,  Kubota M,  Sugimoto T,  Tsusaki K;
XX
DR   WPI; 1995-312772/41.
DR   N-PSDB; AAQ98672.
XX
PT   DNA encoding a trehalose releasing enzyme - which releases trehalose
PT   from a non-reducing saccharide having a trehalose structure as an
PT   end unit.
XX
PS   Claim 6; Page 27-29; 45pp; English.
XX
CC   This enzyme can be used for the preparation of trehalose with high
CC   yields and efficiency from non-reducing saccharides such as
CC   alpha-glucosyltrehalose, alpha-maltosyltrehalose,
CC   alpha-maltotriosyltrehalose, alpha-maltotetrasyltrehalose and
CC   maltopentaosyltrehalose.  The trehalose can be used as a sweetener,
CC   taste-improving agent, quality-improving agent, stabiliser, filler,
CC   excipient or adjuvant in food products cosmetics and pharmaceuticals.
XX
SQ   Sequence   598 AA;
```

```
  Query Match             35.6%;  Score 1060.5;  DB 16;  Length 598;
  Best Local Similarity   39.7%;  Pred. No. 1.6e-78;
  Matches  232;  Conservative   92;  Mismatches  197;  Indels   63;  Gaps   16;

Qy       1 MTFAYKIDGNEVI-----FTLWAPYQKSVKLKVLEKG-LYEMER------DEKGYFTI-- 46
           ||  |  :  : :     : :|||  :|| |  |  |  | |:|      :: |::|
Db       1 mthtypreaakpvlgparydvwapnaesvtl--laggeryamqrraetgpedagwwtaag 58

Qy      47 --TLNNVKVRDRYKYVLD-DASEIPDPASRYQPEGVHGPSQIIQESK-EFNNETFLKKE- 101
             |  ||  .  | |:|| | : :||| :| ||:|||  |:    |   : :: :  :|
Db      59 aptdgnvd----ygylldgdetplpdprtrrqpdgvhalsrtfdpsayswqddawqgrel 114

Qy     102 -DLIIYEIHVGTFTPEGTFEGVIRKLDYLKDLGITAIEIMPIAQFPGKRDWGYDGVYLYA 160
            :|||:|:|||||||| |    |||||  ||:  ||::|:  | |  :||||||  :|
Db     115 qgaviyelhlgtftpegtleaaagkldylaglgvdfiellpvnafngthnwgydgvqwfa 174

Qy     161 VQNSYGGPEGFRKLVDEAHKKGLGVILDVVYNHVGPEGNYMVKLGPYFSQKYKTPWGLTF 220
```

```
          |  :|||||  :::  ||  ||  |||||  ||||||:||  |||: : |||  |     || :
Db    175 vheayggpeayqrfvdaahaaglgviqdvvynhlgpsgnylprfgpylkqgegntwgdsv 234

Qy    221 NFDDAESDEVRKFILENVEYWIKEYNVDGFRLDAVHAIIDTSPKHILEEIADVVHKYNRI 280
          | |   || ||::||:|:  |:::| ||| |||||||: |    ||||:   :  : :
Db    235 nldgpgsdhvrryildnlamwlrdyrvdglrldavhalkderavhiledfgaladqisae 294

Qy    281 V------IAESDLNDPRVVNPKEKCGYNIDAQWVDDFHHSIHAYLTGERQGYYTDFGNLD 334
          |      |||||||:||:: |::  || :: || |||||::|  :|||  |||:|| :|
Db    295 vgrpltliaesdlnnprllyprdvngyglegqwsddfhhavhvnvtgettgyysdfdsla 354

Qy    335 DIVKSYKDVFVYDGKYSNFRRKTHGEPVGELDGCNF--------VVYIQNHDQVGNRGKG 386
          : |  :| | :|| ||:|| : || |:       ||        ||  |||||:|||  |
Db    355 alakvlrdgffhdgsyssfrerhhgrpi------nfsavhpaalvvcsqnhdqignratg 408

Qy    387 ERIIKLVDRESYKIAAALYLLSPYIPMIFMGEEYGEENPFYFFSDFSDSKLIQGVREGRK 446
          :|: : :   |  :|| | |  |: ||: ||||||   |: ||:   : :| :   |||
Db    409 drlsqtlpygslalaavltltgpftpmllmgeeygastpwqfftshpepelgkataegri 468

Qy    447 KE---NGQD----TDPQDESTFNASKLSWKIDEE-----IFSFYKILIKMRKELSIAC-- 492
          ||    | |     ||||  ||  ||| |    |     :   |: |  :|:
Db    469 kefermgwdpavvpdpqdpetfrrskldwaeaaegdharllelyrsltalrrstpdltkl 528

Qy    493 ---DRRVNVVNGENWLIIKGREYFSLYVFSKSSIEVKYSGTLLL 533
             | :|       ||  :       |  ||:  : :  :|| ||
Db    529 gfedtqvafdedarwlrfrrggvqvllnfseqpvsldgagtall 572


RESULT   7
AAR80290
ID   AAR80290 standard; Protein; 597 AA.
XX
AC   AAR80290;
XX
DT   19-JAN-1996  (first entry)
XX
DE   Trehalose releasing enzyme.
XX
KW   Enzyme; trehalose; alpha-glucosyltrehalose; alpha-maltosyltrehalose;
KW   alpha-maltotriosyltrehalose; alpha-maltotetrasyltrehalose;
KW   maltopentaosyltrehalose; sweetener; taste-improving agent;
KW   stabiliser; filler; excipient; adjuvant; foodstuffs; cosmetics;
KW   pharmaceuticals.
XX
OS   Arthrobacter sp. Q36.
XX
PN   EP671470-A2.
XX
PD   13-SEP-1995.
XX
PF   07-MAR-1995;   95EP-0301474.
XX
PR   07-MAR-1994;   94JP-0059840.
PR   07-MAR-1994;   94JP-0059834.
XX
PA   (HAYB ) HAYASHIBARA SEIBUTSU KAGAKU.
XX
PI   Hattori K,  Kubota M,  Sugimoto T,  Tsusaki K;
XX
DR   WPI; 1995-312772/41.
DR   N-PSDB; AAQ98670.
XX
PT   DNA encoding a trehalose releasing enzyme - which releases trehalose
PT   from a non-reducing saccharide having a trehalose structure as an
PT   end unit.
XX
PS   Claim 3; Page 24-25; 45pp; English.
XX
CC   This enzyme can be used for the preparation of trehalose with high
CC   yields and efficiency from non-reducing saccharides such as
CC   alpha-glucosyltrehalose, alpha-maltosyltrehalose,
CC   alpha-maltotriosyltrehalose, alpha-maltotetrasyltrehalose and
CC   maltopentaosyltrehalose.  The trehalose can be used as a sweetener,
CC   taste-improving agent, quality-improving agent, stabiliser, filler,
CC   excipient or adjuvant in food products cosmetics and pharmaceuticals.
XX
SQ   Sequence   597 AA;


  Query Match             35.5%;  Score 1058;  DB 16;  Length 597;
  Best Local Similarity   40.5%;  Pred. No. 2.5e-78;
  Matches  229;  Conservative   88;  Mismatches  191;  Indels   58;  Gaps   15;

Qy     14 FTLWAPYQKSVKLKVLEKG-LYEMER------DEKGYFTI----TLNNVKVRDRYKYVLD 62
          : :|||  :|| |  |  |  | |:|       :: |::|     |  ||     | |:||
Db     18 ydvwapnaesvtl--laggeryamqrraetgpedagwwtaagaptdgnvd----ygylld 71

Qy     63 -DASEIPDPASRYQPEGVHGPSQIIQESK-EFNNETFLKKE--DLIIYEIHVGTFTPEGT 118
           | : :||| :| ||:|||  |:    |   : :: :  :|    :|||:|:|||||||
Db     72 gdetplpdprtrrqpdgvhalsrtfdpsayswqddawggrelqgaviyelhlgtftpegt 131

Qy    119 FEGVIRKLDYLKDLGITAIEIMPIAQFPGKRDWGYDGVYLYAVQNSYGGPEGFRKLVDEA 178
           |    |||||  ||:  ||::|:  | |  :||||||  :||   |||||:::  || |
Db    132 leaaagkldylaglgvdfiellpvnafngthnwgydgvqwfavhedyggpeayqrfvdaa 191

Qy    179 HKKGLGVILDVVYNHVGPEGNYMVKLGPYFSQKYKTPWGLTFNFDDAESDEVRKFILENV 238
          |  ||||| ||||||:|| |||: : |||  |     || : | |   || ||::||:|:
Db    192 haaglgviqdvvynhlgpsgnylprfgpylkqgegntwgdsvnldgpgsdhvrryildnl 251

Qy    239 EYWIKEYNVDGFRLDAVHAIIDTSPKHILEEIADVVHKYNRIV------IAESDLNDPRV 292
            |:::| ||| |||||||: |    ||||:   :  : :  |      |||||||:||:
Db    252 amwlrdyrvdglrldavhalkderavhiledfgaladqisaevgrpltliaesdlnnprl 311

Qy    293 VNPKEKCGYNIDAQWVDDFHHSIHAYLTGERQGYYTDFGNLDDIVKSYKDVFVYDGKYSN 352
          : |::  || :: || ||||||::|  :|||  |||:|| :|  : : |  :| | :|| ||:
Db    312 lyprdvngyglegqwsddfhhavhvnvtgettgyysdfdslaalakvlrdgffhdgsyss 371

Qy    353 FRRKTHGEPVGELDGCNF--------VVYIQNHDQVGNRGKGERIIKLVDRESYKIAAAL 404
          || : || |:      ||        ||  |||||:|||  |:|: : :   |  :|| |
Db    372 frerhhgrpi------nfsavhpaalvvcsqnhdqignratgdrlsqtlpygslalaavl 425

Qy    405 YLLSPYIPMIFMGEEYGEENPFYFFSDFSDSKLIQGVREGRKKE---NGQD----TDPQD 457
           |  |: ||: ||||||   |: ||:   : :| :   ||| ||    | |     ||||
Db    426 tltgpftpmllmgeeygastpwqfftshpepelgkataegrikefermgwdpavvpdpqd 485

Qy    458 ESTFNASKLSWKIDEE-----IFSFYKILIKMRKELSIAC-----DRRVNVVNGENWLII 507
            ||  ||| |    |     :   |: |  :|:           | :|       ||
Db    486 petfrrskldwaeaaegdharllelyrsltalrrstpdltklgfedtqvafdedarwlrf 545

Qy    508 KGREYFSLYVFSKSSIEVKYSGTLLL 533
          :      |  ||:  : :  :|| ||
Db    546 rrggvqvllnfseqpvsldgagtall 571


RESULT   8
AAR80289
ID   AAR80289 standard; Protein; 589 AA.
XX
AC   AAR80289;
XX
DT   19-JAN-1996  (first entry)
XX
DE   Trehalose releasing enzyme.
XX
KW   Enzyme; trehalose; alpha-glucosyltrehalose; alpha-maltosyltrehalose;
KW   alpha-maltotriosyltrehalose; alpha-maltotetrasyltrehalose;
KW   maltopentaosyltrehalose; sweetener; taste-improving agent;
KW   stabiliser; filler; excipient; adjuvant; foodstuffs; cosmetics;
KW   pharmaceuticals.
XX
OS   Rhizobium sp. M11.
XX
PN   EP671470-A2.
XX
PD   13-SEP-1995.
XX
```

PF 07-MAR-1995; 95EP-0301474.
XX
PR 07-MAR-1994; 94JP-0059840.
PR 07-MAR-1994; 94JP-0059834.
XX
PA (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
XX
PI Hattori K, Kubota M, Sugimoto T, Tsusaki K;
XX
DR WPI; 1995-312772/41.
DR N-PSDB; AAQ98669.
XX
PT DNA encoding a trehalose releasing enzyme - which releases trehalose
PT from a non-reducing saccharide having a trehalose structure as an
PT end unit.
XX
PS Claim 3; Page 21-22; 45pp; English.
XX
CC This enzyme can be used for the preparation of trehalose with high
CC yields and efficiency from non-reducing saccharides such as
CC alpha-glucosyltrehalose, alpha-maltosyltrehalose,
CC alpha-maltotriosyltrehalose, alpha-maltotetrasyltrehalose and
CC maltopentaosyltrehalose. The trehalose can be used as a sweetener,
CC taste-improving agent, quality-improving agent, stabiliser, filler,
CC excipient or adjuvant in food products cosmetics and pharmaceuticals.
XX
SQ Sequence 589 AA;

Query Match 34.4%; Score 1025; DB 16; Length 589;
Best Local Similarity 40.3%; Pred. No. 1.3e-75;
Matches 229; Conservative 80; Mismatches 197; Indels 62; Gaps 16;

```
Qy     14 FTLWAPYQKSVKLKVLEKG-LYEMER------DEKGYFTITLNNVKVRDRYKYVLD-DAS 65
          | :|||   :| |  |  |  ||| |       ::|::|           | |:|| |
Db     10 fdiwapeagtvtl--laggeryemgrrpgngpadegwwtaadaptgadvdygylldgdei 67

Qy     66 EIPDPASRYQPEGVHGPSQIIQESKEFNNETFLKKEDL---IIYEIHVGTFTPEGTFEGV 122
          :||| :| ||||||  |:          :  :  :|   :|||:|:|||||||| :
Db     68 plpdprtrrqpegvhalsrtfdpgahrwqdagwqgrelqgsviyelhigtftpegtldaa 127

Qy    123 IRKLDYLKDLGITAIEIMPIAQFPGKRDWGYDGVYLYAVQNSYGGPEGFRKLVDEAHKKG 182
            |||||  |||  ||::|:  | |  :||||||  :||   ||||  ::: || ||  |
Db    128 agkldylaglgidfiellpvnafngthnwgydgvqwfavhegyggpaayqrfvdaahaag 187

Qy    183 LGVILDVVYNHVGPEGNYMVKLGPYFSQKYKTPWGLTFNFDDAESDEVRKFILENVEYWI 242
          |||| ||||||:|| |||: : |||         || : | |   || ||::||:||  |:
Db    188 lgviqdvvynhlgpsgnylprygpylkhgegntwgdsvnldgpgsdhvrqyildnvamwl 247

Qy    243 KEYNVDGFRLDAVHAIIDTSPKHILEE---IADVVHKYN---RIVIAESDLNDPRVVNPK 296
          ::| ||| |||||||: |    |||||   :|| :         :|||||||:||:: |:
Db    248 rdyrvdglrldavhalkderavhileefgaladalsseggrpltliaesdlnnprllypr 307

Qy    297 EKCGYNIDAQWVDDFHHSIHAYLTGERQGYYTDFGNLDDIVKSYKDVFVYDGKYSNFRRK 356
          :  || :  || |||||::|  ::||  |||:|| :|  : |  :| | :|| ||:|| :
Db    308 dvngyglagqwsddfhhavhvnvsgettgyysdfdslgalakvlrdgffhdgsyssfrgr 367

Qy    357 THGEPVGELDGCNF--------VVYIQNHDQVGNRGKGERIIKLVDRESYKIAAALYLLS 408
          || |:      ||         ||  |||||:|||  |:|: : :   |  :|| | |
Db    368 chgrpi------nfsavhpaalvvcsqnhdqignratgdrlsqslpygslalaavltltg 421

Qy    409 PYIPMIFMGEEYGEENPFYFFSDFSDSKLIQGVREGRKKE---NGQD----TDPQDESTF 461
          |: ||:|||||||   |: ||:  : :| :  ||| :|    | |     ||||  ||
Db    422 pftpmlfmgeeygattpwqfftshpepelgkataegrirefermgwdpavvpdpqdpetf 481

Qy    462 NASKLSWK-----IDEEIFSFYKILIKMRK------ELSIACDRRVNVVNGENWLIIKGR 510
            ||| |         :  |: || :|:       | | | |   :   ||
Db    482 trskldwaeasagdharllelyrslitlrrstpelarlgfa-dtavefdddarwl----- 535

Qy    511 EYFS-----LYVFSKSSIEVKYSGTLLL 533
           |:      :  |:   | :   || ||
Db    536 rywrggvqvvlnfadrpisldrpgtall 563
```

RESULT 9
AAR77470
ID AAR77470 standard; Protein; 596 AA.
XX
AC AAR77470;
XX
DT 25-JAN-1996 (first entry)
XX
DE Trehalose releasing enzyme.
XX
KW Enzyme; trehalose; alpha-glucosyltrehalose; alpha-maltosyltrehalose;
KW alpha-maltotriosyltrehalose; alpha-maltotetrasyltrehalose;
KW maltopentaosyltrehalose; sweetener; taste-improving agent;
KW stabiliser; filler; excipient; adjuvant; foodstuffs; cosmetics;
KW pharmaceuticals.
XX
OS Rhizobium sp. M11.
XX
PN EP671470-A2.
XX
PD 13-SEP-1995.
XX
PF 07-MAR-1995; 95EP-0301474.
XX
PR 07-MAR-1994; 94JP-0059840.
PR 07-MAR-1994; 94JP-0059834.
XX
PA (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
XX
PI Hattori K, Kubota M, Sugimoto T, Tsusaki K;
XX
DR WPI; 1995-312772/41.
DR N-PSDB; AAQ98671.
XX
PT DNA encoding a trehalose releasing enzyme - which releases trehalose
PT from a non-reducing saccharide having a trehalose structure as an
PT end unit.
XX
PS Claim 3; Page 27-29; 45pp; English.
XX
CC This enzyme can be used for the preparation of trehalose with high
CC yields and efficiency from non-reducing saccharides such as
CC alpha-glucosyltrehalose, alpha-maltosyltrehalose,
CC alpha-maltotriosyltrehalose, alpha-maltotetrasyltrehalose and
CC maltopentaosyltrehalose. The trehalose can be used as a sweetener,
CC taste-improving agent, quality-improving agent, stabiliser, filler,
CC excipient or adjuvant in food products cosmetics and pharmaceuticals.
XX
SQ Sequence 596 AA;

Query Match 34.4%; Score 1025; DB 16; Length 596;
Best Local Similarity 40.3%; Pred. No. 1.3e-75;
Matches 229; Conservative 80; Mismatches 197; Indels 62; Gaps 16;

```
Qy     14 FTLWAPYQKSVKLKVLEKG-LYEMER------DEKGYFTITLNNVKVRDRYKYVLD-DAS 65
          | :|||   :| |  |  |  ||| |       ::|::|           | |:|| |
Db     17 fdiwapeagtvtl--laggeryemgrrpgngpadegwwtaadaptgadvdygylldgdei 74

Qy     66 EIPDPASRYQPEGVHGPSQIIQESKEFNNETFLKKEDL---IIYEIHVGTFTPEGTFEGV 122
          :||| :| ||||||  |:          :  :  :|   :|||:|:|||||||| :
Db     75 plpdprtrrqpegvhalsrtfdpgahrwqdagwqgrelqgsviyelhigtftpegtldaa 134

Qy    123 IRKLDYLKDLGITAIEIMPIAQFPGKRDWGYDGVYLYAVQNSYGGPEGFRKLVDEAHKKG 182
            |||||  |||  ||::|:  | |  :||||||  :||   ||||  ::: || ||  |
Db    135 agkldylaglgidfiellpvnafngthnwgydgvqwfavhegyggpaayqrfvdaahaag 194

Qy    183 LGVILDVVYNHVGPEGNYMVKLGPYFSQKYKTPWGLTFNFDDAESDEVRKFILENVEYWI 242
          |||| ||||||:|| |||: : |||         || : | |   || ||::||:||  |:
Db    195 lgviqdvvynhlgpsgnylprygpylkhgegntwgdsvnldgpgsdhvrqyildnvamwl 254

Qy    243 KEYNVDGFRLDAVHAIIDTSPKHILEE---IADVVHKYN---RIVIAESDLNDPRVVNPK 296
```

```
          ::| ||| |||||||: |    |||||   :|| :         :|||||||:||:: |:
Db    255 rdyrvdglrldavhalkderavhileefgaladalsseggrpltliaesdlnnprllypr 314

Qy    297 EKCGYNIDAQWVDDFHHSIHAYLTGERQGYYTDFGNLDDIVKSYKDVFVYDGKYSNFRRK 356
          :  || :  || |||||::|  ::||  |||:|| :|  : |  :| | :|| ||:|| :
Db    315 dvngyglagqwsddfhhavhvnvsgettgyysdfdslgalakvlrdgffhdgsyssfrgr 374

Qy    357 THGEPVGELDGCNF--------VVYIQNHDQVGNRGKGERIIKLVDRESYKIAAALYLLS 408
           || |:      ||        ||  |||||:|||  |:|: : :   |  :|| | |
Db    375 chgrpi------nfsavhpaalvvcsqnhdqignratgdrlsqslpygslalaavltltg 428

Qy    409 PYIPMIFMGEEYGEENPFYFFSDFSDSKLIQGVREGRKKE---NGQD----TDPQDESTF 461
          |: ||:|||||||   |: ||:   : :| :   ||| :|    | |     ||||  ||
Db    429 pftpmlfmgeeygattpwqfftshpepelgkataegrirefermgwdpavvpdpqdpetf 488

Qy    462 NASKLSWK-----IDEEIFSFYKILIKMRK------ELSIACDRRVNVVNGENWLIIKGR 510
            ||| |          :   |: || :|:       | | |  |   :   ||
Db    489 trskldwaeasagdharllelyrslitlrrstpelarlgfa-dtavefdddarwl----- 542

Qy    511 EYFS-----LYVFSKSSIEVKYSGTLLL 533
           |:      :  |:   | :   || ||
Db    543 rywrggvqvvlnfadrpisldrpgtall 570


RESULT  10
AAY85157
ID   AAY85157 standard; protein; 575 AA.
XX
AC   AAY85157;
XX
DT   23-JUN-2000  (first entry)
XX
DE   Trehalose-releasing enzyme amino acid sequence.
XX
KW   Non-reducing saccharide-forming enzyme; trehalose; disaccharide; food;
KW   high yield; trehalose releasing enzyme; moisture retention; cosmetic;
KW   production.
XX
OS   Arthrobacter sp. S34.
XX
PN   EP990704-A2.
XX
PD   05-APR-2000.
XX
PF   13-SEP-1999;   99EP-0307220.
XX
PR   11-SEP-1998;   98JP-0258394.
PR   11-DEC-1998;   98JP-0352252.
PR   26-JAN-1999;   99JP-0016931.
XX
PA   (HAYB ) HAYASHIBARA SEIBUTSU KAGAKU.
XX
PI   Yamamoto T,  Maruta K,  Kubota M,  Fukuda S,  Miyake T;
XX
DR   WPI; 2000-273225/24.
DR   N-PSDB; AAA10503.
XX
PT   New mesophilic enzyme from Arthrobacter, used to convert starch
PT   hydrolysate into non-reducing sugars, particularly trehalose, useful as
PT   a sweetener in foods, pharmaceuticals and cosmetics -
XX
PS   Claim 29; Page 47-48; 93pp; English.
XX
CC   This sequence represents a trehalose-releasing enzyme amino acid sequence
CC   from Arthrobacter sp. S34. This enzyme specifically hydrolyses a
CC   non-reducing saccharide having a trehalose structure as an end unit. The
CC   invention relates to this enzyme and also to a non-reducing saccharide
CC   forming enzyme that forms a non-reducing saccharide having a trehalose
CC   structure as an end unit by reducing partial starch hydrolysates. The two
CC   enzymes can be used to produce trehalose, a disaccharide consisting of
CC   two moles of glucose bound at their reducing residues. The disaccharide
CC   is substantially free of reducibility and has moisture retaining
CC   abilities. The enzymes are used to produce non-reducing sugars, they are
CC   particularly used for further enzymatic conversion of starch hydrolysate
CC   to trehalose. The non-reducing sugars produced by the enzymes are useful
CC   in foods, pharmaceuticals  and cosmetics. Non-reducing sugars are
CC   compatible with amino acids and proteins, they do not cause browning, and
CC   have good moisture-retaining properties. The enzymes produce the
CC   non-reducing sugars in high yield and function at a medium temperature
CC   range i.e. between 40 degrees celsius and 60 degrees celsius, and in an
CC   acidic pH range, i.e. a pH of less than 7. Production of non-reducing
CC   sugars using the enzymes eliminates the need to convert a starch
CC   hydrolysate into sugar alcohols (non-reducing) by catalytic
CC   hydrogenation.
XX
SQ   Sequence   575 AA;


  Query Match             34.4%;  Score 1024.5;  DB 21;  Length 575;
  Best Local Similarity   39.4%;  Pred. No. 1.4e-75;
  Matches  215;  Conservative   90;  Mismatches  189;  Indels   51;  Gaps   13;

Qy     14 FTLWAPYQKSVKLKVLEKGLYE--MERDEKGYFTITL---NNVKVRDRYKYVLDDASEIP 68
          | :|||    | | |: :|  |  : ||| |:: :         : | | |::|
Db      5 fpvwapqaaqvtl-vvgqgraelpltrdengwwalqqpwdggpdlvd-ygylvdgkgpfa 62

Qy     69 DPASRYQPEGVHGPSQIIQESKEFNNETFLKKED---------LIIYEIHVGTFTPEGTF 119
          || |  || |||       :  :||:   :   :|          :|||:||||||||||
Db     63 dprslrqprgvh------elgrefdparyawgddgwrgrdltgaviyelhvgtftpegtl 116

Qy    120 EGVIRKLDYLKDLGITAIEIMPIAQFPGKRDWGYDGVYLYAVQNSYGGPEGFRKLVDEAH 179
          :  ||:||:|  ||: |:|::|:  | |   ||||||  |||   ||||| ::: ||  |
Db    117 dsairrldhlvrlgvdavellpvnafngthgwgydgvlwyavhepyggpeayqrfvdach 176

Qy    180 KKGLGVILDVVYNHVGPEGNYMVKLGPYFSQKYKTPWGLTFNFDDAESDEVRKFILENVE 239
           :|| |: ||||||:|| ||::   |||        ||   | |   |||||::|::|
Db    177 arglavvqdvvynhlgpsgnhlpdfgpylgsgaantwgdalnldgplsdevrryiidnav 236

Qy    240 YWIKEYNVDGFRLDAVHAIIDTSPKHILEEIA----DVVHKYNR--IVIAESDLNDPRVV 293
          ||::: : || |||||||: |    |:|||:|    ::  :  |   :|||||||||:::
Db    237 ywlrdmhadglrldavhalrdaralhlleelaarvdelagelgrpltliaesdlndpkli 296

Qy    294 NPKEKCGYNIDAQWVDDFHHSIHAYLTGERQGYYTDFGNLDDIVKSYKDVFVYDGKYSNF 353
            :   || :|||| || ||::|| :|||  ||| ||| |  :|| ::  : :|| :|:|
Db    297 rsraahgygldaqwdddvhhavhanvtgetvgyyadfgglgalvkvfqrgwfhdgtwssf 356

Qy    354 RRKTHGEPVG-ELDGCNFVVYIQNHDQVGNRGKGERIIKLVDRESYKIAAALYLLSPYIP 412
          | : || |:  ::     | : |:|||||||  |:|:   |   |   |||| || |: |
Db    357 rerhhgrpldpdipfrrlvafaqdhdqvgnravgdrmsaqvgegslaaaaalvllgpftp 416

Qy    413 MIFMGEEYGEENPFYFFSDFSDSKLIQGVREGRKKE---NGQD----TDPQDESTFNASK 465
          |:|||||:|   |: ||:   : :| :    ||  |    | |     |||| :||  |
Db    417 mlfmgeewgartpwqfftshpepelgeatargriaefarmgwdpavvpdpqdpatfarsh 476

Qy    466 LSWKIDEE-----IFSFYKILIKMRKELSIACDRRVNVVNGENWLIIKGREYFSLYVFSK 520
          | |   |      : :||  || :|:|| :    |           :   |   :: ||:
Db    477 ldwseperephagllafytdlialrrelpvdapare----------vdadeargvfafsr 526

Qy    521 SSIEV 525
           : |
Db    527 gplrv 531


RESULT  11
AAY85165
ID   AAY85165 standard; Protein; 575 AA.
XX
AC   AAY85165;
XX
DT   23-JUN-2000  (first entry)
XX
DE   Trehalose-releasing enzyme amino acid sequence.
XX
KW   Non-reducing saccharide-forming enzyme; trehalose; disaccharide; food;
KW   high yield; trehalose releasing enzyme; moisture retention; cosmetic;
KW   production.
```

```
XX
OS   Arthrobacter sp. S34.
XX
PN   EP990704-A2.
XX
PD   05-APR-2000.
XX
PF   13-SEP-1999;   99EP-0307220.
XX
PR   11-SEP-1998;   98JP-0258394.
PR   11-DEC-1998;   98JP-0352252.
PR   26-JAN-1999;   99JP-0016931.
XX
PA   (HAYB ) HAYASHIBARA SEIBUTSU KAGAKU.
XX
PI   Yamamoto T,  Maruta K,  Kubota M,  Fukuda S,  Miyake T;
XX
DR   WPI; 2000-273225/24.
DR   N-PSDB; AAA10516.
XX
PT   New mesophilic enzyme from Arthrobacter, used to convert starch
PT   hydrolysate into non-reducing sugars, particularly trehalose, useful as
PT   a sweetener in foods, pharmaceuticals and cosmetics -
XX
PS   Example 6-2; Page 56-58; 93pp; English.
XX
CC   This sequence represents a trehalose-releasing enzyme amino acid sequence
CC   from Arthrobacter sp. S34. This enzyme specifically hydrolyses a
CC   non-reducing saccharide having a trehalose structure as an end unit. The
CC   invention relates to this enzyme and also to a non-reducing saccharide
CC   forming enzyme that forms a non-reducing saccharide having a trehalose
CC   structure as an end unit by reducing partial starch hydrolysates. The two
CC   enzymes can be used to produce trehalose, a disaccharide consisting of
CC   two moles of glucose bound at their reducing residues. The disaccharide
CC   is substantially free of reducibility and has moisture retaining
CC   abilities. The enzymes are used to produce non-reducing sugars, they are
CC   particularly used for further enzymatic conversion of starch hydrolysate
CC   to trehalose. The non-reducing sugars produced by the enzymes are useful
CC   in foods, pharmaceuticals  and cosmetics. Non-reducing sugars are
CC   compatible with amino acids and proteins, they do not cause browning, and
CC   have good moisture-retaining properties. The enzymes produce the
CC   non-reducing sugars in high yield and function at a medium temperature
CC   range i.e. between 40 degrees celsius and 60 degrees celsius, and in an
CC   acidic pH range, i.e. a pH of less than 7. Production of non-reducing
CC   sugars using the enzymes eliminates the need to convert a starch
CC   hydrolysate into sugar alcohols (non-reducing) by catalytic
CC   hydrogenation.
XX
SQ   Sequence   575 AA;


  Query Match             34.4%;  Score 1024.5;  DB 21;  Length 575;
  Best Local Similarity   39.4%;  Pred. No. 1.4e-75;
  Matches  215;  Conservative   90;  Mismatches  189;  Indels   51;  Gaps   13;

Qy      14 FTLWAPYQKSVKLKVLEKGLYE--MERDEKGYFTITL---NNVKVRDRYKYVLDDASEIP 68
           | :|||    | | |: :|  |  : ||| |:: :         : | | |::|
Db       5 fpvwapqaaqvtl-vvgqgraelpltrdengwwalqqpwdgggpdlvd-ygylvdgkgpfa 62

Qy      69 DPASRYQPEGVHGPSQIIQESKEFNNETFLKKED---------LIIYEIHVGTFTPEGTF 119
           || |  || |||      :  :||:   :   :|          :|||:||||||||||
Db      63 dprslrqprgvh------elgrefdparyawgddgwrgrdltgaviyelhvgtftpegtl 116

Qy     120 EGVIRKLDYLKDLGITAIEIMPIAQFPGKRDWGYDGVYLYAVQNSYGGPEGFRKLVDEAH 179
           :  ||:||:|  ||: |:|::|:  | |   ||||||  |||   ||||| ::: ||  |
Db     117 dsairrldhlvrlgvdavellpvnafngthgwgydgvlwyavhepyggpeayqrfvdach 176

Qy     180 KKGLGVILDVVYNHVGPEGNYMVKLGPYFSQKYKTPWGLTFNFDDAESDEVRKFILENVE 239
            :|| |: ||||||:|| ||::   |||        ||   | |   |||||::|:::|
Db     177 arglavvqdvvynhlgpsgnhlpdfgpylgsgaantwgdalnldgplsdevrryiidnav 236

Qy     240 YWIKEYNVDGFRLDAVHAIIDTSPKHILEEIA----DVVHKYNR--IVIAESDLNDPRVV 293
           ||::: : || |||||||: |    |:|||:|    ::  :  |   :|||||||||:::
Db     237 ywlrdmhadglrldavhalrdaralhlleelaarvdelagelgrpltliaesdlndpkli 296

Qy     294 NPKEKCGYNIDAQWVDDFHHSIHAYLTGERQGYYTDFGNLDDIVKSYKDVFVYDGKYSNF 353
             :   || :|||| || ||::|| :|||  ||| ||| |  :|| ::  : :|| :|:|
Db     297 rsraahgygldaqwdddvhhavhanvtgetvgyyadfgglgalvkvfqrgwfhdgtwssf 356

Qy     354 RRKTHGEPVG-ELDGCNFVVYIQNHDQVGNRGKGERIIKLVDRESYKIAAALYLLSPYIP 412
           | : || |:  ::     | : |:|||||||  |:|:   |   |   |||| || |: |
Db     357 rerhhgrpldpdipfrrlvafaqdhdqvgnravgdrmsaqvgegslaaaaalvllgpftp 416

Qy     413 MIFMGEEYGEENPFYFFSDFSDSKLIQGVREGRKKE---NGQD----TDPQDESTFNASK 465
           |:|||||:|   |: ||:   : :| :    ||  |    | |     |||| :||  |
Db     417 mlfmgeewgartpwqfftshpepelgeatargriaefarmgwdpavvpdpqdpatfarsh 476

Qy     466 LSWKIDEE-----IFSFYKILIKMRKELSIACDRRVNVVNGENWLIIKGREYFSLYVFSK 520
           | |   |      : :||  || :|:|| :    |           :   |   :: ||:
Db     477 ldwseperephagllafytdlialrrelpvdapare----------vdadeargvfafsr 526

Qy     521 SSIEV 525
             : |
Db     527 gplrv 531


RESULT  12
AAR80037
ID   AAR80037 standard; Protein; 652 AA.
XX
AC   AAR80037;
XX
DT   11-JUL-1996  (first entry)
XX
DE   Bacillus stearothermophilus thermostable branching enzyme.
XX
KW   Thermostable; 1,4-alpha-glucan; branching enzyme; amylose;
KW   amylopectin; transferase.
XX
OS   Bacillus stearothermophilus (strain TRBE14).
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1
FT                   /note= "Met residue corresponds to TTG codon"
XX
PN   JP07213287-A.
XX
PD   15-AUG-1995.
XX
PF   04-NOV-1993;   93JP-0311004.
XX
PR   04-NOV-1993;   93JP-0311004.
XX
PA   (EZAK ) EZAKI GLICO CO.
XX
DR   WPI; 1995-315926/41.
DR   N-PSDB; AAT04542.
XX
PT   Thermostable branching enzyme gene and recombinant plasmid contg. it
PT    - used to produced a thermostable branching enzyme
XX
PS   Claim 3; Page 6-12; 14pp; Japanese.
XX
CC   A gene coding for a thermostable branching enzyme has been isolated
CC   from chromosomal DNA of B.stearothermophilus strain TRBE14 and
CC   sequenced (see AAT04542). The enzyme (AAR80037) was found to have an
CC   activity of 80% or more at up to 65 deg.C
XX
SQ   Sequence   652 AA;


  Query Match             13.4%;  Score 399;  DB 16;  Length 652;
  Best Local Similarity   25.2%;  Pred. No. 4.2e-24;
  Matches  149;  Conservative   96;  Mismatches  189;  Indels  158;  Gaps   31;

Qy       4 AYKIDGNEVI---------FTLWAPYQKSVKL----KVLEKGLYEMER-DEKGYFTITLN 49
```

```
          :|:: |  ||         | :|||: : |:|       :   : :|: :::| :|| :
Db     22 syelfgahvineggkvgtrfcvwapharevrlvgsfndwdgtdfrlekvndegvwtivvp 81

Qy     50 NVKVRDRYKY--VLDDASEI--PDPASRYQPEGVHGPSQIIQESK--EFNNETFLKK--- 100
             |||  |  |   :   || : |     |  | |  : |  ::|:::: :|
Db     82 enleghlykyeivtpdgqvlfkadpyafyselrphtas-iaydlkgyqwndqswkrkkrr 140

Qy    101 -----EDLIIYEIHVGTF-TPEGTF----EGVIRKLDYLKDLGITAIEIMPIAQFPGKRD 150
              : ::|||:| |::   :| |    |     : |: | | | ||::|: : |  |
Db    141 kriydqpmviyelhfgswkkkdgrfytyremadelisyvldhgfthiellplvehpldrs 200

Qy    151 WGYDGVYLYAVQNSYGGPEGFRKLVDEAHKKGLGVILDVVYNHV--GPEGNYMVKLGP-- 206
          |||  |   ||| : || |  |   ||  |: |:|||:| |  |      | ||    |
Db    201 wgyqgtgyyavtsrygtphdfmyfvdrchqagigvimdwvpghfckdahglymfdgapty 260

Qy    207 -YFSQKYKT--PWGLTFNFDDAESDEVRKFILENVEYWIKEYNVDGFRLDAVHAII---- 259
           | ::| :    || | || |    ||| |:: |  :|:: |::|||:|||  ::
Db    261 eyanekdrenyvwg-tanf-dlgkpevrsflisnalfwleyyhidgfrvdavanmlywpn 318

Qy    260 ------DTSPKHILEEIADVVHKY--NRIVIAESDLNDPRVVNPKEKCGYNIDAQWVDDF 311
                :      | :: : |  |  | ::|||   : |:|  |  : |   : :|
Db    319 ndrlyenpyaveflrklneavfaydpnalmiaedstdwpkvtaptyegglgfnykw---- 374

Qy    312 HHSIHAYLTGERQGYYTDFGNLDDIVKSYKDVFVYDGKYSNFRRKTHGEPVGELDGCNFV 371
                        : | ::|::| | :   |:      ||  | :   |
Db    375 ----------------nmgwmndmlk-ymetppye------rrhvhnqvtfsl----ly 406

Qy    372 VYIQN------HDQVGNRGKGERIIKLVDRESYKIAAALYLLSPYI------PMIFMGEE 419
           | :|      ||:| : ||   : |:      |  | | ||  |:       ::||| |
Db    407 aysenfilpfshdevvh-gkksllnkmpgsyeekf-aqlrllygymmahpgkkllfmgne 464

Qy    420 YGEENPFYFFSDFSDSKLIQGVREGRKKENGQDTDPQDESTFNASKLSWKIDEEI----- 474
                  |: | :                                  || ::|:
Db    465 ---------faqfde---------------------------------wkfedeldwvlf 482

Qy    475 -FSFYKILIKMRKELSIACDRRVNVV-------NGENWLIIKGREYFSLYVF 518
           |  :: :    ||| ||| :|           |  |: :   |  |:: |
Db    483 dfelhrkmndymkel-iacykrykpfyeldhdpqgfewidvhnaeq-sifsf 532


RESULT  13
AAR96109
ID   AAR96109 standard; Protein; 652 AA.
XX
AC   AAR96109;
XX
DT   20-AUG-1996  (first entry)
XX
DE   Starch-branching-enzyme.
XX
KW   Starch-branching-enzyme; thermostable; Q-enzyme; Escherichia coli;
KW   starch; amylopectin; glucan; cyclodextrin; food; food-additive;
KW   adhesive; biodegradable plastic.
XX
OS   Bacillus stearothermophilus strain TRBE14 (FERM P-13916).
XX
FH   Key             Location/Qualifiers
FT   Peptide         238..249
FT                   /note= "Conserved amylase sequence used to
FT                          construct primer AAT27126"
FT   Peptide         309..317
FT                   /note= "Conserved amylase sequence"
FT   Peptide         414..423
FT                   /note= "Conserved amylase sequence used to
FT                          construct primer AAT27127"
XX
PN   EP710674-A2.
XX
PD   08-MAY-1996.
XX
PF   13-SEP-1995;   95EP-0250222.
XX
PR   31-JUL-1995;   95JP-0195647.
PR   13-SEP-1994;   94JP-0218554.
XX
PA   (EZAK ) EZAKI GLICO CO LTD.
XX
PI   Fujii K,  Imanaka T,  Nakamura H,  Okada S,  Takaha T;
PI   Takata H,  Terada Y, Yanase M;
XX
DR   WPI; 1996-223407/23.
DR   N-PSDB; AAT27128.
XX
PT   New glucan with cyclic, branched core and outer branched region -
PT   useful in foods, drinks, adhesives, infusion compsns. etc., readily
PT   soluble in water forming solns. of low viscosity
XX
PS   Example A-2; Page 28-31; 50pp; English.
XX
CC   This sequence represents a thermostable starch- branching-enzyme
CC   (Q-enzyme, EC-2.4.1.18)  from Bacillus stearothermophilus TRBE14
CC   (FERM P-13916).  Conserved regions present in amylases are used to
CC   construct primers AAT27126-7, which are used in PCR to amplify a
CC   fragment of the gene, which is then used as a probe to isolate a
CC   full-length clone in Escherichia coli. The recombinant enzyme may
CC   be reacted with starch or amylopectin to produce a glucan with a
CC   degree of polymerisation of at least 50, with an inner branched
CC   cyclic region and an outer branched region. The product is freely
CC   soluble in water, forming a low-viscosity solution, is more
CC   resistance to degradation than native starch, does not react with
CC   proteins and amino acids, and is degraded to glucose in vivo, so is
CC   easily digestible.  The glucan may be used in foods, beverages,
CC   food-additives, adhesives, in cyclodextrin production and as a
CC   starch substitute in biodegradable plastic.
XX
SQ   Sequence   652 AA;


  Query Match             13.4%;  Score 399;  DB 17;  Length 652;
  Best Local Similarity   25.2%;  Pred. No. 4.2e-24;
  Matches  149;  Conservative   96;  Mismatches  189;  Indels  158;  Gaps   31;

Qy      4 AYKIDGNEVI---------FTLWAPYQKSVKL----KVLEKGLYEMER-DEKGYFTITLN 49
            :|:: |  ||         | :|||: : |:|       :   : :|: :::| :|| :
Db     22 syelfgahvineggkvgtrfcvwapharevrlvgsfndwdgtdfrlekvndegvwtivvp 81

Qy     50 NVKVRDRYKY--VLDDASEI--PDPASRYQPEGVHGPSQIIQESK--EFNNETFLKK--- 100
             |||  |  |   :   || : |     |  | |  : |  ::|:::: :|
Db     82 enleghlykyeivtpdgqvlfkadpyafyselrphtas-iaydlkgyqwndqswkrkkrr 140

Qy    101 -----EDLIIYEIHVGTF-TPEGTF----EGVIRKLDYLKDLGITAIEIMPIAQFPGKRD 150
              : ::|||:| |::   :| |    |     : |: | | | ||::|: : |  |
Db    141 kriydqpmviyelhfgswkkkdgrfytyremadelisyvldhgfthiellplvehpldrs 200

Qy    151 WGYDGVYLYAVQNSYGGPEGFRKLVDEAHKKGLGVILDVVYNHV--GPEGNYMVKLGP-- 206
          |||  |   ||| : || |  |   ||  |: |:|||:| |  |      | ||    |
Db    201 wgyqgtgyyavtsrygtphdfmyfvdrchqagigvimdwvpghfckdahglymfdgapty 260

Qy    207 -YFSQKYKT--PWGLTFNFDDAESDEVRKFILENVEYWIKEYNVDGFRLDAVHAII---- 259
           | ::| :    || | || |    ||| |:: |  :|:: |::|||:|||  ::
Db    261 eyanekdrenyvwg-tanf-dlgkpevrsflisnalfwleyyhidgfrvdavanmlywpn 318

Qy    260 ------DTSPKHILEEIADVVHKY--NRIVIAESDLNDPRVVNPKEKCGYNIDAQWVDDF 311
                :      | :: : |  |  | ::|||   : |:|  |  : |   : :|
Db    319 ndrlyenpyaveflrklneavfaydpnalmiaedstdwpkvtaptyegglgfnykw---- 374

Qy    312 HHSIHAYLTGERQGYYTDFGNLDDIVKSYKDVFVYDGKYSNFRRKTHGEPVGELDGCNFV 371
                        : | ::|::| | :   |:      ||  | :   |
Db    375 ----------------nmgwmndmlk-ymetppye------rrhvhnqvtfsl----ly 406

Qy    372 VYIQN------HDQVGNRGKGERIIKLVDRESYKIAAALYLLSPYI------PMIFMGEE 419
           | :|      ||:| : ||   : |:      |  | | ||  |:       ::||| |
Db    407 aysenfilpfshdevvh-gkksllnkmpgsyeekf-aqlrllygymmahpgkkllfmgne 464

Qy    420 YGEENPFYFFSDFSDSKLIQGVREGRKKENGQDTDPQDESTFNASKLSWKIDEEI----- 474
                  |: | :                                  || ::|:
```

```
Db    465 ---------faqfde---------------------------------wkfedeldwvlf 482

Qy    475 -FSFYKILIKMRKELSIACDRRVNVV-------NGENWLIIKGREYFSLYVF 518
Db    483 dfelhrkmndymkel-iacykrykpfyeldhdpqgfewidvhnaeq-sifsf 532


RESULT  14
AAP94635
ID   AAP94635 standard; protein; 648 AA.
XX
AC   AAP94635;
XX
DT   29-JUN-1990  (first entry)
XX
DE   B. thuringiensis alpha-amylase.
XX
KW   Alpha-amylase; silage; probiotic; lactic acid bacteria;
KW   Lactobacillus plantarum; E. faecalis;
KW   Bacillus thuringiensis strain H1.1 serotype H1 var thuringiensis.
XX
OS   Bacillus thuringiensis.
XX
PN   WO8901970-A.
XX
PD   09-MAR-1989.
XX
PF   02-SEP-1988;   88WO-EP00813.
XX
PR   02-SEP-1987;   87EP-0401972.
XX
PA   (PLAN-) PLANT GENETIC SYST (UYLO-).
XX
PI   Michiels F,  Delcour J,  Mahillon J,  Joos H,  Platteeuw C,  Josson K;
XX
DR   WPI; 1989-085544/11.
DR   N-PSDB; AAN91246.
XX
PT   Transformed lactic acid bacteria -
PT   using exogenous DNA hich codes enzyme that breaks down
PT   oligosaccharide and/or polysaccharide, used in silage or as a
PT   probiotic
XX
PS   Fig 2; pages 2/35-3/35; 112pp; English.
XX
CC   The inventors claim an inoculum for silage which comprises lactic
CC   acid bacteria, esp. Lactobacillus plantarum and E. faecalis, transformed
CC   with an exogenous DNA encoding an enzyme which breaks down an oligosacc-
CC   -haride and/or a polysaccharide into a fermentable carbohydrate. The
CC   exogenous enzyme is especially an alpha-amylase (AAP94635) encoded by
CC   the DNA sequence in AAN91246 (claims 2, 6 and 10). The transformed lactic
CC   acid bacteria can break down polysaccharides and oligosaccharides in a
CC   silage crop to provide a complete source of the fermentable carbohydrates
CC   that the lactic acid bacteria need for producing lactic acid. Also
CC   claimed is a probiotic characterised by the transformed lactic acid
CC   bacteria which can be used for establishng and maintaining optimal
CC   intestinal flora in animals, including humans.
XX
SQ   Sequence   648 AA;


  Query Match             12.9%;  Score 385.5;  DB 10;  Length 648;
  Best Local Similarity   23.7%;  Pred. No. 5.3e-23;
  Matches  149;  Conservative  117;  Mismatches  231;  Indels  131;  Gaps   29;

Qy      4 AYKIDGNEVI---------FTLWAPYQKSVK-LKVLEKGLYEMER----DEKGYFTITLN 49
Db     25 synifgahvvtedeiqgvrftvwaphakamsvvgdfnewdyeqhkmlqvteegiwslfip 84

Qy     50 NVKVRDRYKYVLD----DASEIPDPASRYQPEGVHGPSQIIQ-ESKEFNNETF------- 97
Db     85 hieegeiykyaietlagdvilkadpyavyaevrpntasvvfdikgyewndknwnrkkkkk 144

Qy     98 -LKKEDLIIYEIHVGTF--TPEGTF----EGVIRKLDYLKDLGITAIEIMPIAQFPGKRD 150
Db    145 siykeamtvyelhfgswkkkedgtlysyremveelipyvvehqfthieimplvehpydrs 204

Qy    151 WGYDGVYLYAVQNSYGGPEGFRKLVDEAHKKGLGVILDVVYNHV--GPEGNYMVKLGPYF 208
Db    205 wgyqgtgyyaatsrfgtphdlmhfvdechkygigvildwvpghfckdahglylfdgtpty 264

Qy    209 SQKYK-----TPWGLTFNFDDAESDEVRKFILENVEYWIKEYNVDGFRLDAVHAII---- 259
Db    265 qykdkdvqenpvwg-tvnfdlgkr-evrnflisnalfwmryfhidgfrvdavanmlywnk 322

Qy    260 ---DTSPKH---ILEEIADVVHKYNR--IVIAESDLNDPRVVNPKEKCGYNIDAQWVDDF 311
Db    323 egqeqsnehavsflrelneavfaededflmtaedstawplvttptyegglgfnykw---- 378

Qy    312 HHSIHAYLTGERQGYYTDFGNLDDIVKSYKDVFVYDGKYSNFRRKTHGEPVGEL---DGC 368
Db    379 -----------------nmgwmndvlk-------ymecapeyrkhihekmtfsllyayse 414

Qy    369 NFVVYIQNHDQVGNRGKGERIIKLVDRESYKIAAALYLLSPYI------PMIFMGEEYGE 422
Db    415 nfilpl-shdevvh-gk-ksllnkmpgdywdkfaqlrllygyffthpgekllfmggefgq 471

Qy    423 ENPFYFFSDFSD-SKLIQGVREGRKKENGQDTDPQDESTFNASKLSWKID---------- 471
Db    472 ------fdewkdledldwnlhdfemhrymhdyfnelialykrskplwqldhspegfqwid 525

Qy    472 -----EEIFSFYKILIKMRKELSIACDRRVNVVNGENWLI-IKGREYFSLYVFSKSSIEV 525
Db    526 annneqsifsfirqgdkqedalvvvcn--ftkatyenykvgvpdfeyyneilnsdae--- 580

Qy    526 KYSGT---------LLLSSNNSFPQHIE 544
Db    581 qyggsgqvnkkrlktilepyhnqaahve 608


RESULT  15
AAR11271
ID   AAR11271 standard; Protein; 639 AA.
XX
AC   AAR11271;
XX
DT   29-MAY-1991  (first entry)
XX
DE   B.stearothermophilus branching enzyme.
XX
KW   Branching enzyme; amylaceous substances; transferase;
KW   thermostability.
XX
OS   Bacillus stearothermophilus.
XX
PN   EP418945-A.
XX
PD   27-MAR-1991.
XX
PF   22-AUG-1990;   90EP-0202264.
XX
PR   23-AUG-1989;   89NL-0002128.
XX
PA   (COOP-) COOP VERKOOP AVEBE.
XX
PI   Kiel JAKW,  Boels JM,  Beldman G,  Venema G;
XX
DR   WPI; 1991-088587/13.
DR   N-PSDB; AAQ11072.
XX
PT   Branching enzyme obtd. from Bacillus stearothermophilus - has good
PT   thermostability, for use in modifying starch-like substances
XX
PS   Disclosure; fig 4; 13pp; English.
XX
CC   This branching enzyme is derived from B.stearothermophilus and is
```

CC encoded by the glgB gene. It is stable at high temps. and can be
CC produced in large amts. using recombinant techniques. It is a
CC transferase allowing formation of alpha(1-6) branches in starch-like
CC substances. It is useful for extending the shelf life of such
CC materials and can be used in the prepn. of human and animal foods.
XX
SQ Sequence 639 AA;

Query Match 12.6%; Score 374.5; DB 12; Length 639;
Best Local Similarity 23.6%; Pred. No. 4.2e-22;
Matches 141; Conservative 99; Mismatches 190; Indels 167; Gaps 29;

```
Qy       4 AYKIDGNEVI---------FTLWAPYQKSVKL---------------KVLEKGLYEMERD 39
Db      22 syelfgahvikkngmvgtrfcvwapharevrlvgsfnewngtnfnlmkvsnqgvwmifip 81

Qy      40 E--KGYFTITLNNVKVRDRYKYVL--DDASEI--PDPASRYQPEGVHGPSQIIQ-ESKEF 92
Db      82 enleghl-----------ykyeittndgnvllksdpyafyselrphtasivynikgyqw 129

Qy      93 NNETFLKK--------EDLIIYEIHVGTF--TPEGTF----EGVIRKLDYLKDLGITAIE 138
Db     130 ndqtwrrkkqrkriydqplfiyelhfgswkkkedgsfytyqemaeelipyvlehgfthie 189

Qy     139 IMPIAQFPGKRDWGYDGVYLYAVQNSYGGPEGFRKLVDEAHKKGLGVILDVVYNHV--GP 196
Db     190 llplvehpfdrswgyqgigyysatsrygtphdlmyfidrchqagigvildwvpghfckds 249

Qy     197 EGNYMVKLGPYF-----SQKYKTPWGLTFNFDDAESDEVRKFILENVEYWIKEYNVDGFR 251
Db     250 hglymfdgapayeyanmqdrenyvwg-tanf-dlgkpevrsflisnalfwmeyfhvdgfr 307

Qy     252 LDAVHAII----------DTSPKHILEEIADVVHKY--NRIVIAESDLNDPRVVNPKEKC 299
Db     308 vdavanmlywpnsdvlykntyaveflqklnetvfaydpnilmiaedstdwprvtaptydg 367

Qy     300 GYNIDAQWVDDFHHSIHAYLTGERQGYYTDFGNLDDIVKSYKDVFVYDGKYSNFRRKTHG 359
Db     368 glgfnykw--------------------nmgwmndil-------tymetppehrkyvhn 399

Qy     360 EPVGELDGCNFVVYIQN------HDQVGNRGKGERIIKLVDRESYKIAAAL----YLLS- 408
Db     400 kvtfsl----lyaysenfilpfshdevvh-gkksllskmpgtyeekfaqlrllygyllth 454

Qy     409 PYIPMIFMGEEYGEENPFYFFSDFSDSKLIQGVREGRKKENGQDTDPQDESTFNASKLSW 468
Db     455 pgkkllfmggefgq------fdewkd---------------------------leqldw 480

Qy     469 KIDEEIFSFYKILIKMRKELSIACDRRVNVV-------NGENWLIIKGREYFSLYVF 518
Db     481 mlfd--fdmhrnmnmyvkel-lkcykrykplyeldhspdgfewidvhnaeq-sifsf 533
```

Search completed: July 15, 2002, 12:17:32
Job time: 69 sec

GenCore version 4.5

OM protein - protein search, using sw model

Run on: July 15, 2002, 12:16:28 ; Search time 25.09 Seconds
(without alignments)
544.197 Million cell updates/sec

Title: US-09-298-924-6
Perfect score: 2978
Sequence: 1 MTFAYKIDGNEVIFTLWAPY.....PQHIEEGKYEFDKGFALYKL 559

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2978	100.0	559	4	US-09-242-690A-15	Sequence 15, Appl
2	1748.5	58.7	556	2	US-08-505-377-1	Sequence 1, Appli
3	1748.5	58.7	556	3	US-08-798-269-1	Sequence 1, Appli
4	1748.5	58.7	556	4	US-09-055-210-1	Sequence 1, Appli
5	1060.5	35.6	598	1	US-08-399-646-14	Sequence 14, Appl
6	1060.5	35.6	598	1	US-08-607-321-14	Sequence 14, Appl
7	1060.5	35.6	598	2	US-08-961-240-14	Sequence 14, Appl
8	1060.5	35.6	598	2	US-08-605-501-14	Sequence 14, Appl
9	1059	35.6	597	1	US-08-399-646-4	Sequence 4, Appli
10	1059	35.6	597	1	US-08-607-321-4	Sequence 4, Appli
11	1059	35.6	597	2	US-08-961-240-4	Sequence 4, Appli
12	1059	35.6	597	2	US-08-605-501-4	Sequence 4, Appli
13	1025	34.4	589	1	US-08-399-646-2	Sequence 2, Appli
14	1025	34.4	589	1	US-08-607-321-2	Sequence 2, Appli
15	1025	34.4	589	2	US-08-961-240-2	Sequence 2, Appli
16	1025	34.4	589	2	US-08-605-501-2	Sequence 2, Appli
17	1025	34.4	596	1	US-08-399-646-12	Sequence 12, Appl
18	1025	34.4	596	1	US-08-607-321-12	Sequence 12, Appl
19	1025	34.4	596	2	US-08-961-240-12	Sequence 12, Appl
20	1025	34.4	596	2	US-08-605-501-12	Sequence 12, Appl
21	399	13.4	652	4	US-08-528-026C-4	Sequence 4, Appli
22	346.5	11.6	713	4	US-09-346-237-9	Sequence 9, Appli
23	343	11.5	718	4	US-09-346-237-10	Sequence 10, Appl
24	334	11.2	606	4	US-09-187-124-2	Sequence 2, Appli
25	331.5	11.1	893	4	US-09-514-302-4	Sequence 4, Appli
26	331.5	11.1	1938	4	US-09-514-302-2	Sequence 2, Appli
27	329.5	11.1	818	2	US-08-410-784A-2	Sequence 2, Appli
28	329.5	11.1	818	4	US-09-346-237-11	Sequence 11, Appl
29	321.5	10.8	740	2	US-08-410-784A-4	Sequence 4, Appli
30	303.5	10.2	829	4	US-09-514-599-6	Sequence 6, Appli
31	303.5	10.2	862	4	US-09-346-237-1	Sequence 1, Appli
32	295	9.9	555	1	US-08-039-777-3	Sequence 3, Appli
33	293.5	9.9	555	1	US-08-611-361A-3	Sequence 3, Appli
34	293.5	9.9	921	4	US-09-514-599-2	Sequence 2, Appli
35	291	9.8	915	4	US-09-346-237-2	Sequence 2, Appli
36	291	9.8	928	4	US-09-514-599-4	Sequence 4, Appli
37	286	9.6	928	1	US-08-474-140-11	Sequence 11, Appl
38	286	9.6	928	1	US-08-477-630-11	Sequence 11, Appl
39	286	9.6	928	1	US-08-472-293-11	Sequence 11, Appl
40	286	9.6	928	1	US-08-474-545-11	Sequence 11, Appl
41	286	9.6	928	2	US-08-478-341-11	Sequence 11, Appl
42	286	9.6	928	3	US-08-996-733-11	Sequence 11, Appl
43	285	9.6	963	1	US-08-537-002A-3	Sequence 3, Appli
44	285	9.6	963	3	US-08-863-010-3	Sequence 3, Appli
45	285	9.6	963	4	US-09-024-429-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-09-242-690A-15
; Sequence 15, Application US/09242690A
; Patent No. 6284534
; GENERAL INFORMATION:
; APPLICANT: KONDO, KEIJI
; APPLICANT: MIURA, YUTAKA
; TITLE OF INVENTION: YEAST VECTOR AND METHOD OF PRODUCING PROTEINS USING THE
; TITLE OF INVENTION: SAME
; FILE REFERENCE: 049441/0118
; CURRENT APPLICATION NUMBER: US/09/242,690A
; CURRENT FILING DATE: 1999-02-23
; PRIOR APPLICATION NUMBER: PCT/JP97/02924
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: JP 8/241062
; PRIOR FILING DATE: 1996-08-23
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 559
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic protein
US-09-242-690A-15

102(e) date for this patent is 2/23/99
Junior to mine

Query Match 100.0%; Score 2978; DB 4; Length 559;
Best Local Similarity 100.0%; Pred. No. 4.1e-250;
Matches 559; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 MTFAYKIDGNEVIFTLWAPYQKSVKLKVLEKGLYEMERDEKGYFTITLNNVKVRDRYKYV 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MTFAYKIDGNEVIFTLWAPYQKSVKLKVLEKGLYEMERDEKGYFTITLNNVKVRDRYKYV 60

Qy      61 LDDASEIPDPASRYQPEGVHGPSQIIQESKEFNNETFLKKEDLIIYEIHVGTFTPEGTFE 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 LDDASEIPDPASRYQPEGVHGPSQIIQESKEFNNETFLKKEDLIIYEIHVGTFTPEGTFE 120

Qy     121 GVIRKLDYLKDLGITAIEIMPIAQFPGKRDWGYDGVYLYAVQNSYGGPEGFRKLVDEAHK 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 GVIRKLDYLKDLGITAIEIMPIAQFPGKRDWGYDGVYLYAVQNSYGGPEGFRKLVDEAHK 180

Qy     181 KGLGVILDVVYNHVGPEGNYMVKLGPYFSQKYKTPWGLTFNFDDAESDEVRKFILENVEY 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 KGLGVILDVVYNHVGPEGNYMVKLGPYFSQKYKTPWGLTFNFDDAESDEVRKFILENVEY 240

Qy     241 WIKEYNVDGFRLDAVHAIIDTSPKHILEEIADVVHKYNRIVIAESDLNDPRVVNPKEKCG 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 WIKEYNVDGFRLDAVHAIIDTSPKHILEEIADVVHKYNRIVIAESDLNDPRVVNPKEKCG 300
```

```
Qy    301 YNIDAQWVDDFHHSIHAYLTGERQGYYTDFGNLDDIVKSYKDVFVYDGKYSNFRRKTHGE 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 YNIDAQWVDDFHHSIHAYLTGERQGYYTDFGNLDDIVKSYKDVFVYDGKYSNFRRKTHGE 360

Qy    361 PVGELDGCNFVVYIQNHDQVGNRGKGERIIKLVDRESYKIAAALYLLSPYIPMIFMGEEY 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 PVGELDGCNFVVYIQNHDQVGNRGKGERIIKLVDRESYKIAAALYLLSPYIPMIFMGEEY 420

Qy    421 GEENPFYFFSDFSDSKLIQGVREGRKKENGQDTDPQDESTFNASKLSWKIDEEIFSFYKI 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GEENPFYFFSDFSDSKLIQGVREGRKKENGQDTDPQDESTFNASKLSWKIDEEIFSFYKI 480

Qy    481 LIKMRKELSIACDRRVNVVNGENWLIIKGREYFSLYVFSKSSIEVKYSGTLLLSSNNSFP 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 LIKMRKELSIACDRRVNVVNGENWLIIKGREYFSLYVFSKSSIEVKYSGTLLLSSNNSFP 540

Qy    541 QHIEEGKYEFDKGFALYKL 559
          |||||||||||||||||||
Db    541 QHIEEGKYEFDKGFALYKL 559
```

```
RESULT   2
US-08-505-377-1
; Sequence 1, Application US/08505377
; Patent No. 5856146
;  GENERAL INFORMATION:
;    APPLICANT:  MITSUZUMI, Hitoshi
;    APPLICANT:  KUBOTA, Michio
;    APPLICANT:  SUGIMOTO, Toshiyuki
;    TITLE OF INVENTION:  RECOMBINANT THERMOSTABLE ENZYME WHICH
;    TITLE OF INVENTION:  RELEASES TREHALOSE FROM NON-REDUCING SACCHARIDE
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Browdy and Neimark
;      STREET:  419 Seventh Street N.W. Ste. 300
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/505,377
;      FILING DATE:  21-JUL-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 190180/1994
;      FILING DATE:  21-JUL-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 109128/1995
;      FILING DATE:  11-APR-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP NOT YET RECEIVED
;      FILING DATE:  04-JUL-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Browdy, Roger L
;      REGISTRATION NUMBER:  25,618
;      REFERENCE/DOCKET NUMBER:  MITSUZUMI=1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 628-5197
;      TELEFAX:  (202) 737-3528
;      TELEX:  249688
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  556 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-505-377-1
```

```
  Query Match           58.7%;  Score 1748.5;  DB 2;  Length 556;
  Best Local Similarity  58.9%;  Pred. No. 1.9e-143;
  Matches  329;  Conservative   87;  Mismatches  136;  Indels    7;  Gaps    5;

Qy      2 TFAYKIDGNEVIFTLWAPYQKSVKLKVLEKGLYEMERDEKGYFTITLNNVKVRDRYKYVL 61
            :|   |: |: || |||||  ||||| | | |  ||::::|:| : :::::    | |::
Db      3 SFGGNIEKNKGIFKLWAPYVNSVKLK-LSKKLIPMEKNDEGFFEVEIDDIEENLTYSYII 61

Qy     62 DDASEIPDPASRYQPEGVHGPSQIIQESKEFNNETFLKKEDLIIYEIHVGTFTPEGTFEG 121
            :|  |||||||||| ||| ||:|:   :  :   :| |||||||:|||||: || |:|
Db     62 EDKREIPDPASRYQPLGVHDKSQLIRTDYQILDLGKVKIEDLIIYELHVGTFSQEGNFKG 121

Qy    122 VIRKLDYLKDLGITAIEIMPIAQFPGKRDWGYDGVYLYAVQNSYGGPEGFRKLVDEAHKK 181
            || |||||||||| ||:||:|||| |||||||:||||||:|||    |||:||||:
Db    122 VIEKLDYLKDLGITGIELMPVAQFPGNRDWGYDGVFLYAVQNTYGGPWELAKLVNEAHKR 181

Qy    182 GLGVILDVVYNHVGPEGNYMVKLGPYFSQKYKTPWGLTFNFDDAESDEVRKFILENVEYW 241
            |: |||||||||:||||||:: |||||| :|||||||||||||    |:||||||||||
Db    182 GIAVILDVVYNHIGPEGNYLLGLGPYFSDRYKTPWGLTFNFDDRGCDQVRKFILENVEYW 241

Qy    242 IKEYNVDGFRLDAVHAIIDTSPKHILEEIADVVHKYNRIVIAESDLNDPRVVNPKEKCGY 301
            | : :|| |||||||| | ||||||:|||:  |:  : |||||||||::|  |: |||
Db    242 FKTFKIDGLRLDAVHAIFDNSPKHILQEIAEKAHQLGKFVIAESDLNDPKIV--KDDCGY 299

Qy    302 NIDAQWVDDFHHSIHAYLTGERQGYYTDFGNLDDIVKSYKDVFVYDGKYSNFRRKTHGEP 361
            ||||||||||::||::| |:  || ||| ::|| |::||||||||||| :| :||| |
Db    300 KIDAQWVDDFHHAVHAFITKEKDYYYQDFGRIEDIEKTFKDVFVYDGKYSRYRGRTHGAP 359

Qy    362 VGELDGCNFVVYIQNHDQVGNRGKGERIIKLVDRESYKIAAALYLLSPYIPMIFMGEEYG 421
            ||:|    |||:||||||||||| |||:  | |: :| :|| ||:||||||:|||||||
Db    360 VGDLPPRKFVVFIQNHDQVGNRGNGERLSILTDKTTYLMAATLYILSPYIPLIFMGEEYY 419

Qy    422 EENPFYFFSDFSDSKLIQGVREGRKKENGQDTDPQDESTFNASKLSWKIDEEIFSFYKIL 481
            | |||:|||||||  ||:|||||| ||| |  ||| |  |  |  |||||||||:  :|| |
Db    420 ETNPFFFFSDFSDPVLIKGVREGRLKENNQMIDPQSEEAFLKSKLSWKIDEEVLDYYKQL 479

Qy    482 IKMRKELSIACDRRVNVVNGENWLIIKGREYFSLYVFSKSSIEVKYSGTLLLSSNNSFPQ 541
            | :||  :  | |  |    | : :  :  :  |   |  | :| ||:    ||:
Db    480 INIRKRYN-NCKRVKEVRREGNCITLIMEKIGIIASFDDIVINSKITGNLLIGI--GFPK 536

Qy    542 HIEEGK-YEFDKGFALYKL 559
            ::: :  : ::|  :|:|
Db    537 KLKKDELIKVNRGVGVYQL 555
```

```
RESULT   3
US-08-798-269-1
; Sequence 1, Application US/08798269
; Patent No. 6027918
;  GENERAL INFORMATION:
;    APPLICANT:  MITSUZUMI, Hitoshi
;    APPLICANT:  KUBOTA, Michio
;    APPLICANT:  SUGIMOTO, Toshiyuki
;    TITLE OF INVENTION:  RECOMBINANT THERMOSTABLE ENZYME WHICH
;    TITLE OF INVENTION:  RELEASES TREHALOSE FROM NON-REDUCING SACCHARIDE
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Browdy and Neimark
;      STREET:  419 Seventh Street N.W. Ste. 300
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/798,269
;      FILING DATE:
```

```
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/505,377
;      FILING DATE:  21-JUL-1995
;      APPLICATION NUMBER:  JP 190180/1994
;      FILING DATE:  21-JUL-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 109128/1995
;      FILING DATE:  11-APR-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP NOT YET RECEIVED
;      FILING DATE:  04-JUL-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Browdy, Roger L
;      REGISTRATION NUMBER:  25,618
;      REFERENCE/DOCKET NUMBER:  MITSUZUMI=1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 628-5197
;      TELEFAX:  (202) 737-3528
;      TELEX:  249688
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  556 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-798-269-1


  Query Match           58.7%;  Score 1748.5;  DB 3;  Length 556;
  Best Local Similarity  58.9%;  Pred. No. 1.9e-143;
  Matches  329;  Conservative   87;  Mismatches  136;  Indels    7;  Gaps    5;

Qy       2 TFAYKIDGNEVIFTLWAPYQKSVKLKVLEKGLYEMERDEKGYFTITLNNVKVRDRYKYVL 61
Db       3 SFGGNIEKNKGIFKLWAPYVNSVKLK-LSKKLIPMEKNDEGFFEVEIDDIEENLTYSYII 61

Qy      62 DDASEIPDPASRYQPEGVHGPSQIIQESKEFNNETFLKKEDLIIYEIHVGTFTPEGTFEG 121
Db      62 EDKREIPDPASRYQPLGVHDKSQLIRTDYQILDLGKVKIEDLIIYELHVGTFSQEGNFKG 121

Qy     122 VIRKLDYLKDLGITAIEIMPIAQFPGKRDWGYDGVYLYAVQNSYGGPEGFRKLVDEAHKK 181
Db     122 VIEKLDYLKDLGITGIELMPVAQFPGNRDWGYDGVFLYAVQNTYGGPWELAKLVNEAHKR 181

Qy     182 GLGVILDVVYNHVGPEGNYMVKLGPYFSQKYKTPWGLTFNFDDAESDEVRKFILENVEYW 241
Db     182 GIAVILDVVYNHIGPEGNYLLGLGPYFSDRYKTPWGLTFNFDDRGCDQVRKFILENVEYW 241

Qy     242 IKEYNVDGFRLDAVHAIIDTSPKHILEEIADVVHKYNRIVIAESDLNDPRVVNPKEKCGY 301
Db     242 FKTFKIDGLRLDAVHAIFDNSPKHILQEIAEKAHQLGKFVIAESDLNDPKIV--KDDCGY 299

Qy     302 NIDAQWVDDFHHSIHAYLTGERQGYYTDFGNLDDIVKSYKDVFVYDGKYSNFRRKTHGEP 361
Db     300 KIDAQWVDDFHHAVHAFITKEKDYYYQDFGRIEDIEKTFKDVFVYDGKYSRYRGRTHGAP 359

Qy     362 VGELDGCNFVVYIQNHDQVGNRGKGERIIKLVDRESYKIAAALYLLSPYIPMIFMGEEYG 421
Db     360 VGDLPPRKFVVFIQNHDQVGNRGNGERLSILTDKTTYLMAATLYILSPYIPLIFMGEEYY 419

Qy     422 EENPFYFFSDFSDSKLIQGVREGRKKENGQDTDPQDESTFNASKLSWKIDEEIFSFYKIL 481
Db     420 ETNPFFFFSDFSDPVLIKGVREGRLKENNQMIDPQSEEAFLKSKLSWKIDEEVLDYYKQL 479

Qy     482 IKMRKELSIACDRRVNVVNGENWLIIKGREYFSLYVFSKSSIEVKYSGTLLLSSNNSFPQ 541
Db     480 INIRKRYN-NCKRVKEVRREGNCITLIMEKIGIIASFDDIVINSKITGNLLIGI--GFPK 536

Qy     542 HIEEGK-YEFDKGFALYKL 559
Db     537 KLKKDELIKVNRGVGVYQL 555
```

```
RESULT   4
US-09-055-210-1
; Sequence 1, Application US/09055210
; Patent No. 6346394
;  GENERAL INFORMATION:
;    APPLICANT:  MITSUZUMI, Hitoshi
;    APPLICANT:  KUBOTA, Michio
;    APPLICANT:  SUGIMOTO, Toshiyuki
;    TITLE OF INVENTION:  RECOMBINANT THERMOSTABLE ENZYME WHICH
;    TITLE OF INVENTION:  RELEASES TREHALOSE FROM NON-REDUCING SACCHARIDE
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Browdy and Neimark
;      STREET:  419 Seventh Street N.W. Ste. 300
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/055,210
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/505,377
;      FILING DATE:  21-JUL-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 109128/1995
;      FILING DATE:  11-APR-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP NOT YET RECEIVED
;      FILING DATE:  04-JUL-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Browdy, Roger L
;      REGISTRATION NUMBER:  25,618
;      REFERENCE/DOCKET NUMBER:  MITSUZUMI=1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 628-5197
;      TELEFAX:  (202) 737-3528
;      TELEX:  249688
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  556 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-055-210-1


  Query Match           58.7%;  Score 1748.5;  DB 4;  Length 556;
  Best Local Similarity  58.9%;  Pred. No. 1.9e-143;
  Matches  329;  Conservative   87;  Mismatches  136;  Indels    7;  Gaps    5;

Qy       2 TFAYKIDGNEVIFTLWAPYQKSVKLKVLEKGLYEMERDEKGYFTITLNNVKVRDRYKYVL 61
Db       3 SFGGNIEKNKGIFKLWAPYVNSVKLK-LSKKLIPMEKNDEGFFEVEIDDIEENLTYSYII 61

Qy      62 DDASEIPDPASRYQPEGVHGPSQIIQESKEFNNETFLKKEDLIIYEIHVGTFTPEGTFEG 121
Db      62 EDKREIPDPASRYQPLGVHDKSQLIRTDYQILDLGKVKIEDLIIYELHVGTFSQEGNFKG 121

Qy     122 VIRKLDYLKDLGITAIEIMPIAQFPGKRDWGYDGVYLYAVQNSYGGPEGFRKLVDEAHKK 181
Db     122 VIEKLDYLKDLGITGIELMPVAQFPGNRDWGYDGVFLYAVQNTYGGPWELAKLVNEAHKR 181

Qy     182 GLGVILDVVYNHVGPEGNYMVKLGPYFSQKYKTPWGLTFNFDDAESDEVRKFILENVEYW 241
Db     182 GIAVILDVVYNHIGPEGNYLLGLGPYFSDRYKTPWGLTFNFDDRGCDQVRKFILENVEYW 241
```

```
Qy    242 IKEYNVDGFRLDAVHAIIDTSPKHILEEIADVVHKYNRIVIAESDLNDPRVVNPKEKCGY 301
          | : :|| |||||||| | ||||||:|||:  |:  : ||||||||||::|  |: |||
Db    242 FKTFKIDGLRLDAVHAIFDNSPKHILQEIAEKAHQLGKFVIAESDLNDPKIV--KDDCGY 299

Qy    302 NIDAQWVDDFHHSIHAYLTGERQGYYTDFGNLDDIVKSYKDVFVYDGKYSNFRRKTHGEP 361
          |||||||||||::||::| |:  || ||| ::|| |::||||||||||| :| :||| |
Db    300 KIDAQWVDDFHHAVHAFITKEKDYYYQDFGRIEDIEKTFKDVFVYDGKYSRYRGRTHGAP 359

Qy    362 VGELDGCNFVVYIQNHDQVGNRGKGERIIKLVDRESYKIAAALYLLSPYIPMIFMGEEYG 421
          ||:|    |||:|||||||||||| |||:  | |: :| :|| ||:|||||:|||||||
Db    360 VGDLPPRKFVVFIQNHDQVGNRGNGERLSILTDKTTYLMAATLYILSPYIPLIFMGEEYY 419

Qy    422 EENPFYFFSDFSDSKLIQGVREGRKKENGQDTDPQDESTFNASKLSWKIDEEIFSFYKIL 481
          | |||:|||||||  ||:|||||| ||| |  ||| |  |  ||||||||||:  :|| |
Db    420 ETNPFFFFSDFSDPVLIKGVREGRLKENNQMIDPQSEEAFLKSKLSWKIDEEVLDYYKQL 479

Qy    482 IKMRKELSIACDRRVNVVNGENWLIIKGREYFSLYVFSKSSIEVKYSGTLLLSSNNSFPQ 541
          | :||  :  | |   |    | : :   :   :  |    |  | :| ||:     ||:
Db    480 INIRKRYN-NCKRVKEVRREGNCITLIMEKIGIIASFDDIVINSKITGNLLIGI--GFPK 536

Qy    542 HIEEGK-YEFDKGFALYKL 559
          ::: :  : ::|  :|:|
Db    537 KLKKDELIKVNRGVGVYQL 555
```

```
RESULT   5
US-08-399-646-14
; Sequence 14, Application US/08399646
; Patent No. 5556781
;  GENERAL INFORMATION:
;    APPLICANT:  KUBOTA, Michio
;    APPLICANT:  TSUSAKI, Keiji
;    APPLICANT:  HATTORI, Kazuko
;    APPLICANT:  SUGIMOTO, Toshiyuki
;    TITLE OF INVENTION:  DNA ENCODING ENZYME, RECOMBINANT DNA AND
;    TITLE OF INVENTION:  ENZYME, TRANSFORMANT, AND THEIR PREPARATIONS AND USE
;    NUMBER OF SEQUENCES:  18
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BROWDY AND NEIMARK
;      STREET:  419 Seventh Street, N.W., Suite 300
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/399,646
;      FILING DATE:  07-MAR-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 59834
;      FILING DATE:  07-MAR-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 59840
;      FILING DATE:  07-MAR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BROWDY, Roger L.
;      REGISTRATION NUMBER:  25,618
;      REFERENCE/DOCKET NUMBER:  KUBOTA=5
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-628-5197
;      TELEFAX:  202-737-3528
;      TELEX:  248633
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  598 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-399-646-14


  Query Match             35.6%;  Score 1060.5;  DB 1;  Length 598;
  Best Local Similarity   39.7%;  Pred. No. 1e-83;
  Matches  232;  Conservative   92;  Mismatches  197;  Indels   63;  Gaps   16;

Qy      1 MTFAYKIDGNEVI-----FTLWAPYQKSVKLKVLEKG-LYEMER------DEKGYFTI-- 46
            ||  |  :  : :      : :|||  :|| |  |  |  | |:|      :: |::|
Db      1 MTHTYPREAAKPVLGPARYDVWAPNAESVTL--LAGGERYAMQRRAETGPEDAGWWTAAG 58

Qy     47 --TLNNVKVRDRYKYVLD-DASEIPDPASRYQPEGVHGPSQIIQESK-EFNNETFLKKE- 101
            |  ||     | |:|| | : :||| :| ||:|||  |:    |   : :: :  :|
Db     59 APTDGNVD----YGYLLDGDETPLPDPRTRRQPDGVHALSRTFDPSAYSWQDDAWQGREL 114

Qy    102 -DLIIYEIHVGTFTPEGTFEGVIRKLDYLKDLGITAIEIMPIAQFPGKRDWGYDGVYLYA 160
             :|||:|:|||||||| |    |||||  ||:  ||::|:  | |  :||||||  :|
Db    115 QGAVIYELHLGTFTPEGTLEAAAGKLDYLAGLGVDFIELLPVNAFNGTHNWGYDGVQWFA 174

Qy    161 VQNSYGGPEGFRKLVDEAHKKGLGVILDVVYNHVGPEGNYMVKLGPYFSQKYKTPWGLTF 220
            |  :|||||  ::: || ||  |||||  ||||||:|| |||: : |||  |     || :
Db    175 VHEAYGGPEAYQRFVDAAHAAGLGVIQDVVYNHLGPSGNYLPRFGPYLKQGEGNTWGDSV 234

Qy    221 NFDDAESDEVRKFILENVEYWIKEYNVDGFRLDAVHAIIDTSPKHILEEIADVVHKYNRI 280
            | |   || ||::||:|:  |:::| ||| |||||||: |    ||||:   :  : :
Db    235 NLDGPGSDHVRRYILDNLAMWLRDYRVDGLRLDAVHALKDERAVHILEDFGALADQISAE 294

Qy    281 V------IAESDLNDPRVVNPKEKCGYNIDAQWVDDFHHSIHAYLTGERQGYYTDFGNLD 334
            |      |||||||:||:: |::  || :: || |||||::|  :|||  |||:|| :|
Db    295 VGRPLTLIAESDLNNPRLLYPRDVNGYGLEGQWSDDFHHAVHVNVTGETTGYYSDFDSLA 354

Qy    335 DIVKSYKDVFVYDGKYSNFRRKTHGEPVGELDGCNF--------VVYIQNHDQVGNRGKG 386
            : |  :| | :|| ||:|| : || |:      ||         ||  |||||:|||  | .
Db    355 ALAKVLRDGFFHDGSYSSFRERHHGRPI------NFSAVHPAALVVCSQNHDQIGNRATG 408

Qy    387 ERIIKLVDRESYKIAAALYLLSPYIPMIFMGEEYGEENPFYFFSDFSDSKLIQGVREGRK 446
            :|: : :   |  :|| | |  |: ||: ||||||   |: ||:   : :| :   |||
Db    409 DRLSQTLPYGSLALAAVLTLTGPFTPMLLMGEEYGASTPWQFFTSHPEPELGKATAEGRI 468

Qy    447 KE---NGQD----TDPQDESTFNASKLSWKIDEE-----IFSFYKILIKMRKELSIAC-- 492
            ||    | |     ||||  ||  ||| |    |     :   |: |  :|:
Db    469 KEFERMGWDPAVVPDPQDPETFRRSKLDWAEAAEGDHARLLELYRSLTALRRSTPDLTKL 528

Qy    493 ---DRRVNVVNGENWLIIKGREYFSLYVFSKSSIEVKYSGTLLL 533
              | :|       ||  :      |  ||:  : :  :|| ||
Db    529 GFEDTQVAFDEDARWLRFRRGGVQVLLNFSEQPVSLDGAGTALL 572
```

```
RESULT   6
US-08-607-321-14
; Sequence 14, Application US/08607321
; Patent No. 5716813
;  GENERAL INFORMATION:
;    APPLICANT:  KUBOTA, Michio
;    APPLICANT:  TSUSAKI, Keiji
;    APPLICANT:  HATTORI, Kazuko
;    APPLICANT:  SUGIMOTO, Toshiyuki
;    TITLE OF INVENTION:  DNA ENCODING ENZYME, RECOMBINANT DNA AND
;    TITLE OF INVENTION:  ENZYME, TRANSFORMANT, AND THEIR PREPARATIONS AND USE
;    NUMBER OF SEQUENCES:  18
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BROWDY AND NEIMARK
;      STREET:  419 Seventh Street, N.W., Suite 300
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
```

```
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/607,321
;      FILING DATE:  26-FEB-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/399,646
;      FILING DATE:  07-MAR-1995
;      APPLICATION NUMBER:  JP 59834
;      FILING DATE:  07-MAR-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 59840
;      FILING DATE:  07-MAR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BROWDY, Roger L.
;      REGISTRATION NUMBER:  25,618
;      REFERENCE/DOCKET NUMBER:  KUBOTA=5
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-628-5197
;      TELEFAX:  202-737-3528
;      TELEX:  248633
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  598 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-607-321-14


  Query Match             35.6%;  Score 1060.5;  DB 1;  Length 598;
  Best Local Similarity   39.7%;  Pred. No. 1e-83;
  Matches  232;  Conservative   92;  Mismatches  197;  Indels   63;  Gaps   16;

Qy       1 MTFAYKIDGNEVI-----FTLWAPYQKSVKLKVLEKG-LYEMER------DEKGYFTI-- 46
Db       1 MTHTYPREAAKPVLGPARYDVWAPNAESVTL--LAGGERYAMQRRAETGPEDAGWWTAAG 58

Qy      47 --TLNNVKVRDRYKYVLD-DASEIPDPASRYQPEGVHGPSQIIQESK-EFNNETFLKKE- 101
Db      59 APTDGNVD----YGYLLDGDETPLPDPRTRRQPDGVHALSRTFDPSAYSWQDDAWQGREL 114

Qy     102 -DLIIYEIHVGTFTPEGTFEGVIRKLDYLKDLGITAIEIMPIAQFPGKRDWGYDGVYLYA 160
Db     115 QGAVIYELHLGTFTPEGTLEAAAGKLDYLAGLGVDFIELLPVNAFNGTHNWGYDGVQWFA 174

Qy     161 VQNSYGGPEGFRKLVDEAHKKGLGVILDVVYNHVGPEGNYMVKLGPYFSQKYKTPWGLTF 220
Db     175 VHEAYGGPEAYQRFVDAAHAAGLGVIQDVVYNHLGPSGNYLPRFGPYLKQGEGNTWGDSV 234

Qy     221 NFDDAESDEVRKFILENVEYWIKEYNVDGFRLDAVHAIIDTSPKHILEEIADVVHKYNRI 280
Db     235 NLDGPGSDHVRRYILDNLAMWLRDYRVDGLRLDAVHALKDERAVHILEDFGALADQISAE 294

Qy     281 V------IAESDLNDPRVVNPKEKCGYNIDAQWVDDFHHSIHAYLTGERQGYYTDFGNLD 334
Db     295 VGRPLTLIAESDLNNPRLLYPRDVNGYGLEGQWSDDFHHAVHVNVTGETTGYYSDFDSLA 354

Qy     335 DIVKSYKDVFVYDGKYSNFRRKTHGEPVGELDGCNF--------VVYIQNHDQVGNRGKG 386
Db     355 ALAKVLRDGFFHDGSYSSFRERHHGRPI------NFSAVHPAALVVCSQNHDQIGNRATG 408

Qy     387 ERIIKLVDRESYKIAAALYLLSPYIPMIFMGEEYGEENPFYFFSDFSDSKLIQGVREGRK 446
Db     409 DRLSQTLPYGSLALAAVLTLTGPFTPMLLMGEEYGASTPWQFFTSHPEPELGKATAEGRI 468

Qy     447 KE---NGQD----TDPQDESTFNASKLSWKIDEE-----IFSFYKILIKMRKELSIAC-- 492
Db     469 KEFERMGWDPAVVPDPQDPETFRRSKLDWAEAAEGDHARLLELYRSLTALRRSTPDLTKL 528

Qy     493 ---DRRVNVVNGENWLIIKGREYFSLYVFSKSSIEVKYSGTLLL 533
Db     529 GFEDTQVAFDEDARWLRFRRGGVQVLLNFSEQPVSLDGAGTALL 572
```

```
RESULT   7
US-08-961-240-14
; Sequence 14, Application US/08961240
; Patent No. 5830715
;  GENERAL INFORMATION:
;    APPLICANT:  KUBOTA, Michio
;    APPLICANT:  TSUSAKI, Keiji
;    APPLICANT:  HATTORI, Kazuko
;    APPLICANT:  SUGIMOTO, Toshiyuki
;    TITLE OF INVENTION:  DNA ENCODING ENZYME, RECOMBINANT DNA AND
;    TITLE OF INVENTION:  ENZYME, TRANSFORMANT, AND THEIR PREPARATIONS AND USE
;    NUMBER OF SEQUENCES:  18
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BROWDY AND NEIMARK
;      STREET:  419 Seventh Street, N.W., Suite 300
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/961,240
;      FILING DATE:  30-OCT-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/399,646
;      FILING DATE:  07-MAR-1995
;      APPLICATION NUMBER:  JP 59834
;      FILING DATE:  07-MAR-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 59840
;      FILING DATE:  07-MAR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BROWDY, Roger L.
;      REGISTRATION NUMBER:  25,618
;      REFERENCE/DOCKET NUMBER:  KUBOTA=5
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-628-5197
;      TELEFAX:  202-737-3528
;      TELEX:  248633
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  598 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-961-240-14


  Query Match             35.6%;  Score 1060.5;  DB 2;  Length 598;
  Best Local Similarity   39.7%;  Pred. No. 1e-83;
  Matches  232;  Conservative   92;  Mismatches  197;  Indels   63;  Gaps   16;

Qy       1 MTFAYKIDGNEVI-----FTLWAPYQKSVKLKVLEKG-LYEMER------DEKGYFTI-- 46
Db       1 MTHTYPREAAKPVLGPARYDVWAPNAESVTL--LAGGERYAMQRRAETGPEDAGWWTAAG 58

Qy      47 --TLNNVKVRDRYKYVLD-DASEIPDPASRYQPEGVHGPSQIIQESK-EFNNETFLKKE- 101
Db      59 APTDGNVD----YGYLLDGDETPLPDPRTRRQPDGVHALSRTFDPSAYSWQDDAWQGREL 114

Qy     102 -DLIIYEIHVGTFTPEGTFEGVIRKLDYLKDLGITAIEIMPIAQFPGKRDWGYDGVYLYA 160
Db     115 QGAVIYELHLGTFTPEGTLEAAAGKLDYLAGLGVDFIELLPVNAFNGTHNWGYDGVQWFA 174

Qy     161 VQNSYGGPEGFRKLVDEAHKKGLGVILDVVYNHVGPEGNYMVKLGPYFSQKYKTPWGLTF 220
Db     175 VHEAYGGPEAYQRFVDAAHAAGLGVIQDVVYNHLGPSGNYLPRFGPYLKQGEGNTWGDSV 234
```

```
Qy    221 NFDDAESDEVRKFILENVEYWIKEYNVDGFRLDAVHAIIDTSPKHILEEIADVVHKYNRI 280
          | |   || ||::||:|:  |:::| ||| |||||||: |    ||||:   :  : :
Db    235 NLDGPGSDHVRRYILDNLAMWLRDYRVDGLRLDAVHALKDERAVHILEDFGALADQISAE 294

Qy    281 V----~~-IAESDLNDPRVVNPKEKCGYNIDAQWVDDFHHSIHAYLTGERQGYYTDFGNLD 334
          |      |||||||:||:: |::  || :: || |||||::|  :||| |||:|| :|
Db    295 VGRPLTLIAESDLNNPRLLYPRDVNGYGLEGQWSDDFHHAVHVNVTGETTGYYSDFDSLA 354

Qy    335 DIVKSYKDVFVYDGKYSNFRRKTHGEPVGELDGCNF--------VVYIQNHDQVGNRGKG 386
           : |  :| | :|| ||:|| : || |:       ||        ||  |||||:|||  |
Db    355 ALAKVLRDGFFHDGSYSSFRERHHGRPI------NFSAVHPAALVVCSQNHDQIGNRATG 408

Qy    387 ERIIKLVDRESYKIAAALYLLSPYIPMIFMGEEYGEENPFYFFSDFSDSKLIQGVREGRK 446
          :|: : :   |  :|| | |  |: ||: ||||||   |: ||:   : :| :   |||
Db    409 DRLSQTLPYGSLALAAVLTLTGPFTPMLLMGEEYGASTPWQFFTSHPEPELGKATAEGRI 468

Qy    447 KE---NGQD----TDPQDESTFNASKLSWKIDEE-----IFSFYKILIKMRKELSIAC-- 492
          ||    | |     |||| ||  ||| |    |     :   |: | :|:
Db    469 KEFERMGWDPAVVPDPQDPETFRRSKLDWAEAAEGDHARLLELYRSLTALRRSTPDLTKL 528

Qy    493 ---DRRVNVVNGENWLIIKGREYFSLYVFSKSSIEVKYSGTLLL 533
             | :|       ||  :      |  ||:  : :  :|| ||
Db    529 GFEDTQVAFDEDARWLRFRRGGVQVLLNFSEQPVSLDGAGTALL 572


RESULT   8
US-08-605-501-14
; Sequence 14, Application US/08605501
; Patent No. 5834287
;  GENERAL INFORMATION:
;    APPLICANT:  KUBOTA, Michio
;    APPLICANT:  TSUSAKI, Keiji
;    APPLICANT:  HATTORI, Kazuko
;    APPLICANT:  SUGIMOTO, Toshiyuki
;    TITLE OF INVENTION:  DNA ENCODING ENZYME, RECOMBINANT DNA AND
;    TITLE OF INVENTION:  ENZYME, TRANSFORMANT, AND THEIR PREPARATIONS AND USE
;    NUMBER OF SEQUENCES:  18
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BROWDY AND NEIMARK
;      STREET:  419 Seventh Street, N.W., Suite 300
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/605,501
;      FILING DATE:  26-FEB-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/399,646
;      FILING DATE:  07-MAR-1995
;      APPLICATION NUMBER:  JP 59834
;      FILING DATE:  07-MAR-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 59840
;      FILING DATE:  07-MAR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BROWDY, Roger L.
;      REGISTRATION NUMBER:  25,618
;      REFERENCE/DOCKET NUMBER:  KUBOTA=5
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-628-5197
;      TELEFAX:  202-737-3528
;      TELEX:  248633
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  598 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-605-501-14


  Query Match             35.6%;  Score 1060.5;  DB 2;  Length 598;
  Best Local Similarity   39.7%;  Pred. No. 1e-83;
  Matches  232;  Conservative   92;  Mismatches  197;  Indels   63;  Gaps   16;

Qy      1 MTFAYKIDGNEVI-----FTLWAPYQKSVKLKVLEKG-LYEMER------DEKGYFTI-- 46
            ||  |  :  : :      : :|||  :|| |  |  |  | |:|       :: |::|
Db      1 MTHTYPREAAKPVLGPARYDVWAPNAESVTL--LAGGERYAMQRRAETGPEDAGWWTAAG 58

Qy     47 --TLNNVKVRDRYKYVLD-DASEIPDPASRYQPEGVHGPSQIIQESK-EFNNETFLKKE- 101
            |  ||      | |:|| | : :||| :| ||:|||  |:     |   : :: :  :|
Db     59 APTDGNVD----YGYLLDGDETPLPDPRTRRQPDGVHALSRTFDPSAYSWQDDAWQGREL 114

Qy    102 -DLIIYEIHVGTFTPEGTFEGVIRKLDYLKDLGITAIEIMPIAQFPGKRDWGYDGVYLYA 160
             :|||:|:|||||||| |    |||||  ||:  ||::|:  | |  :||||||  :|
Db    115 QGAVIYELHLGTFTPEGTLEAAAGKLDYLAGLGVDFIELLPVNAFNGTHNWGYDGVQWFA 174

Qy    161 VQNSYGGPEGFRKLVDEAHKKGLGVILDVVYNHVGPEGNYMVKLGPYFSQKYKTPWGLTF 220
            |  :|||||  ::: || ||  ||||| |||||:|| |||: : |||  |     || :
Db    175 VHEAYGGPEAYQRFVDAAHAAGLGVIQDVVYNHLGPSGNYLPRFGPYLKQGEGNTWGDSV 234

Qy    221 NFDDAESDEVRKFILENVEYWIKEYNVDGFRLDAVHAIIDTSPKHILEEIADVVHKYNRI 280
          | |   || ||::||:|:  |:::| ||| |||||||: |    ||||:   :  : :
Db    235 NLDGPGSDHVRRYILDNLAMWLRDYRVDGLRLDAVHALKDERAVHILEDFGALADQISAE 294

Qy    281 V------IAESDLNDPRVVNPKEKCGYNIDAQWVDDFHHSIHAYLTGERQGYYTDFGNLD 334
          |      |||||||:||:: |::  || :: || |||||::|  :||| |||:|| :|
Db    295 VGRPLTLIAESDLNNPRLLYPRDVNGYGLEGQWSDDFHHAVHVNVTGETTGYYSDFDSLA 354

Qy    335 DIVKSYKDVFVYDGKYSNFRRKTHGEPVGELDGCNF--------VVYIQNHDQVGNRGKG 386
           : |  :| | :|| ||:|| : || |:       ||        ||  |||||:|||  |
Db    355 ALAKVLRDGFFHDGSYSSFRERHHGRPI------NFSAVHPAALVVCSQNHDQIGNRATG 408

Qy    387 ERIIKLVDRESYKIAAALYLLSPYIPMIFMGEEYGEENPFYFFSDFSDSKLIQGVREGRK 446
          :|: : :   |  :|| | |  |: ||: ||||||   |: ||:   : :| :   |||
Db    409 DRLSQTLPYGSLALAAVLTLTGPFTPMLLMGEEYGASTPWQFFTSHPEPELGKATAEGRI 468

Qy    447 KE---NGQD----TDPQDESTFNASKLSWKIDEE-----IFSFYKILIKMRKELSIAC-- 492
          ||    | |     |||| ||  ||| |    |     :   |: | :|:
Db    469 KEFERMGWDPAVVPDPQDPETFRRSKLDWAEAAEGDHARLLELYRSLTALRRSTPDLTKL 528

Qy    493 ---DRRVNVVNGENWLIIKGREYFSLYVFSKSSIEVKYSGTLLL 533
             | :|       ||  :      |  ||:  : :  :|| ||
Db    529 GFEDTQVAFDEDARWLRFRRGGVQVLLNFSEQPVSLDGAGTALL 572


RESULT   9
US-08-399-646-4
; Sequence 4, Application US/08399646
; Patent No. 5556781
;  GENERAL INFORMATION:
;    APPLICANT:  KUBOTA, Michio
;    APPLICANT:  TSUSAKI, Keiji
;    APPLICANT:  HATTORI, Kazuko
;    APPLICANT:  SUGIMOTO, Toshiyuki
;    TITLE OF INVENTION:  DNA ENCODING ENZYME, RECOMBINANT DNA AND
;    TITLE OF INVENTION:  ENZYME, TRANSFORMANT, AND THEIR PREPARATIONS AND USE
;    NUMBER OF SEQUENCES:  18
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BROWDY AND NEIMARK
;      STREET:  419 Seventh Street, N.W., Suite 300
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
```

```
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/399,646
;      FILING DATE:  07-MAR-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 59834
;      FILING DATE:  07-MAR-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 59840
;      FILING DATE:  07-MAR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BROWDY, Roger L.
;      REGISTRATION NUMBER:  25,618
;      REFERENCE/DOCKET NUMBER:  KUBOTA=5
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-628-5197
;      TELEFAX:  202-737-3528
;      TELEX:  248633
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  597 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-399-646-4


  Query Match             35.6%;  Score 1059;  DB 1;  Length 597;
  Best Local Similarity   40.5%;  Pred. No. 1.4e-83;
  Matches  229;  Conservative   89;  Mismatches  190;  Indels   58;  Gaps   15;

Qy      14 FTLWAPYQKSVKLKVLEKG-LYEMER------DEKGYFTI----TLNNVKVRDRYKYVLD 62
           : :|||  :|| |  |  |  | |:|      :: |::|     |  ||     | |:||
Db      18 YDVWAPNAESVTL--LAGGERYAMQRRAETGPEDAGWWTAAGAPTDGNVD----YGYLLD 71

Qy      63 -DASEIPDPASRYQPEGVHGPSQIIQESK-EFNNETFLKKE--DLIIYEIHVGTFTPEGT 118
            | : :||| :| ||:|||  |:    |   : :: :  :|    :|||:|:||||||||
Db      72 GDETPLPDPRTRRQPDGVHALSRTFDPSAYSWQDDAWQGRELQGAVIYELHLGTFTPEGT 131

Qy     119 FEGVIRKLDYLKDLGITAIEIMPIAQFPGKRDWGYDGVYLYAVQNSYGGPEGFRKLVDEA 178
            |    |||||  ||:  ||::|:  | |  :||||||  :||  :|||||  ::: || |
Db     132 LEAAAGKLDYLAGLGVDFIELLPVNAFNGTHNWGYDGVQWFAVHEAYGGPEAYQRFVDAA 191

Qy     179 HKKGLGVILDVVYNHVGPEGNYMVKLGPYFSQKYKTPWGLTFNFDDAESDEVRKFILENV 238
            |  ||||| ||||||:|| |||: : |||  |     || : | |   || ||::||:|:
Db     192 HAAGLGVIQDVVYNHLGPSGNYLPRFGPYLKQGEGNTWGDSVNLDGPGSDHVRRYILDNL 251

Qy     239 EYWIKEYNVDGFRLDAVHAIIDTSPKHILEEIADVVHKYNRIV------IAESDLNDPRV 292
             |:::| ||| |||||||: |    ||||:   :  : :  |      |||||||:||:
Db     252 AMWLRDYRVDGLRLDAVHALKDERAVHILEDFGALADQISAEVGRPLTLIAESDLNNPRL 311

Qy     293 VNPKEKCGYNIDAQWVDDFHHSIHAYLTGERQGYYTDFGNLDDIVKSYKDVFVYDGKYSN 352
           : |::  || :: || |||||::|  :|||  |||:|| :|  : |  :| | :|| ||:
Db     312 LYPRDVNGYGLEGQWSDDFHHAVHVNVTGETTGYYSDFDSLAALAKVLRDGFFHDGSYSS 371

Qy     353 FRRKTHGEPVGELDGCNF--------VVYIQNHDQVGNRGKGERIIKLVDRESYKIAAAL 404
           || : || |:      ||         ||  |||||:|||  |:|: : :   |  :|| |
Db     372 FRERHHGRPI------NFSAVHPAALVVCSQNHDQIGNRATGDRLSQTLPYGSLALAAVL 425

Qy     405 YLLSPYIPMIFMGEEYGEENPFYFFSDFSDSKLIQGVREGRKKE---NGQD----TDPQD 457
            |  |: ||: ||||||   |: ||:   : :| :   ||| ||    | |     ||||
Db     426 TLTGPFTPMLLMGEEYGASTPWQFFTSHPEPELGKATAEGRIKEFERMGWDPAVVPDPQD 485

Qy     458 ESTFNASKLSWKIDEE-----IFSFYKILIKMRKELSIAC-----DRRVNVVNGENWLII 507
           ||  ||| |    |     :   |: | :|:          | :|       ||
Db     486 PETFRRSKLDWAEAAEGDHARLLELYRSLTALRRSTPDLTKLGFEDTQVAFDEDARWLRF 545

Qy     508 KGREYFSLYVFSKSSIEVKYSGTLLL 533
           :      |  ||:  : :  :|| ||
Db     546 RRGGVQVLLNFSEQPVSLDGAGTALL 571
```

```
RESULT  10
US-08-607-321-4
; Sequence 4, Application US/08607321
; Patent No. 5716813
;  GENERAL INFORMATION:
;    APPLICANT:  KUBOTA, Michio
;    APPLICANT:  TSUSAKI, Keiji
;    APPLICANT:  HATTORI, Kazuko
;    APPLICANT:  SUGIMOTO, Toshiyuki
;    TITLE OF INVENTION:  DNA ENCODING ENZYME, RECOMBINANT DNA AND
;    TITLE OF INVENTION:  ENZYME, TRANSFORMANT, AND THEIR PREPARATIONS AND USE
;    NUMBER OF SEQUENCES:  18
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BROWDY AND NEIMARK
;      STREET:  419 Seventh Street, N.W., Suite 300
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/607,321
;      FILING DATE:  26-FEB-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/399,646
;      FILING DATE:  07-MAR-1995
;      APPLICATION NUMBER:  JP 59834
;      FILING DATE:  07-MAR-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 59840
;      FILING DATE:  07-MAR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BROWDY, Roger L.
;      REGISTRATION NUMBER:  25,618
;      REFERENCE/DOCKET NUMBER:  KUBOTA=5
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-628-5197
;      TELEFAX:  202-737-3528
;      TELEX:  248633
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  597 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-607-321-4


  Query Match             35.6%;  Score 1059;  DB 1;  Length 597;
  Best Local Similarity   40.5%;  Pred. No. 1.4e-83;
  Matches  229;  Conservative   89;  Mismatches  190;  Indels   58;  Gaps   15;

Qy      14 FTLWAPYQKSVKLKVLEKG-LYEMER------DEKGYFTI----TLNNVKVRDRYKYVLD 62
           : :|||  :|| |  |  |  | |:|      :: |::|     |  ||     | |:||
Db      18 YDVWAPNAESVTL--LAGGERYAMQRRAETGPEDAGWWTAAGAPTDGNVD----YGYLLD 71

Qy      63 -DASEIPDPASRYQPEGVHGPSQIIQESK-EFNNETFLKKE--DLIIYEIHVGTFTPEGT 118
            | : :||| :| ||:|||  |:    |   : :: :  :|    :|||:|:||||||||
Db      72 GDETPLPDPRTRRQPDGVHALSRTFDPSAYSWQDDAWQGRELQGAVIYELHLGTFTPEGT 131

Qy     119 FEGVIRKLDYLKDLGITAIEIMPIAQFPGKRDWGYDGVYLYAVQNSYGGPEGFRKLVDEA 178
            |    |||||  ||:  ||::|:  | |  :||||||  :||  :|||||  ::: || |
Db     132 LEAAAGKLDYLAGLGVDFIELLPVNAFNGTHNWGYDGVQWFAVHEAYGGPEAYQRFVDAA 191

Qy     179 HKKGLGVILDVVYNHVGPEGNYMVKLGPYFSQKYKTPWGLTFNFDDAESDEVRKFILENV 238
            |  ||||| ||||||:|| |||: : |||  |     || : | |   || ||::||:|:
Db     192 HAAGLGVIQDVVYNHLGPSGNYLPRFGPYLKQGEGNTWGDSVNLDGPGSDHVRRYILDNL 251
```

```
Qy    239 EYWIKEYNVDGFRLDAVHAIIDTSPKHILEEIADVVHKYNRIV------IAESDLNDPRV 292
          |:::| ||| |||||||: |    ||||:   :  : :  |      |||||||:||:
Db    252 AMWLRDYRVDGLRLDAVHALKDERAVHILEDFGALADQISAEVGRPLTLIAESDLNNPRL 311

Qy    293 VNPKEKCGYNIDAQWVDDFHHSIHAYLTGERQGYYTDFGNLDDIVKSYKDVFVYDGKYSN 352
          : |::  || :: || |||||::|  :||| |||:|| :|  : |  :| | :|| ||:
Db    312 LYPRDVNGYGLEGQWSDDFHHAVHVNVTGETTGYYSDFDSLAALAKVLRDGFFHDGSYSS 371

Qy    353 FRRKTHGEPVGELDGCNF--------VVYIQNHDQVGNRGKGERIIKLVDRESYKIAAAL 404
          || : || |:      ||         ||  |||||:|||  |:|: : :   |  :|| |
Db    372 FRERHHGRPI------NFSAVHPAALVVCSQNHDQIGNRATGDRLSQTLPYGSLALAAVL 425

Qy    405 YLLSPYIPMIFMGEEYGEENPFYFFSDFSDSKLIQGVREGRKKE---NGQD----TDPQD 457
          |  |: ||: ||||||   |: ||:   : :| :   ||| ||    | |     ||||
Db    426 TLTGPFTPMLLMGEEYGASTPWQFFTSHPEPELGKATAEGRIKEFERMGWDPAVVPDPQD 485

Qy    458 ESTFNASKLSWKIDEE-----IFSFYKILIKMRKELSIAC-----DRRVNVVNGENWLII 507
          || ||| |    |     :   |: |  :|:          | :|       ||
Db    486 PETFRRSKLDWAEAAEGDHARLLELYRSLTALRRSTPDLTKLGFEDTQVAFDEDARWLRF 545

Qy    508 KGREYFSLYVFSKSSIEVKYSGTLLL 533
          :      |  ||:  : :  :|| ||
Db    546 RRGGVQVLLNFSEQPVSLDGAGTALL 571


RESULT  11
US-08-961-240-4
; Sequence 4, Application US/08961240
; Patent No. 5830715
;  GENERAL INFORMATION:
;    APPLICANT:  KUBOTA, Michio
;    APPLICANT:  TSUSAKI, Keiji
;    APPLICANT:  HATTORI, Kazuko
;    APPLICANT:  SUGIMOTO, Toshiyuki
;    TITLE OF INVENTION:  DNA ENCODING ENZYME, RECOMBINANT DNA AND
;    TITLE OF INVENTION:  ENZYME, TRANSFORMANT, AND THEIR PREPARATIONS AND USE
;    NUMBER OF SEQUENCES:  18
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BROWDY AND NEIMARK
;      STREET:  419 Seventh Street, N.W., Suite 300
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/961,240
;      FILING DATE:  30-OCT-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/399,646
;      FILING DATE:  07-MAR-1995
;      APPLICATION NUMBER:  JP 59834
;      FILING DATE:  07-MAR-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 59840
;      FILING DATE:  07-MAR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BROWDY, Roger L.
;      REGISTRATION NUMBER:  25,618
;      REFERENCE/DOCKET NUMBER:  KUBOTA=5
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-628-5197
;      TELEFAX:  202-737-3528
;      TELEX:  248633
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  597 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-961-240-4


  Query Match             35.6%;  Score 1059;  DB 2;  Length 597;
  Best Local Similarity   40.5%;  Pred. No. 1.4e-83;
  Matches  229;  Conservative   89;  Mismatches  190;  Indels   58;  Gaps   15;

Qy     14 FTLWAPYQKSVKLKVLEKG-LYEMER------DEKGYFTI----TLNNVKVRDRYKYVLD 62
          : :|||  :|| |  |  |  | |:|       :: |::|     |  ||     | |:||
Db     18 YDVWAPNAESVTL--LAGGERYAMQRRAETGPEDAGWWTAAGAPTDGNVD----YGYLLD 71

Qy     63 -DASEIPDPASRYQPEGVHGPSQIIQESK-EFNNETFLKKE--DLIIYEIHVGTFTPEGT 118
           | : :||| :| ||:||| |:    |   : :: :  :|    :|||:|:|||||||
Db     72 GDETPLPDPRTRRQPDGVHALSRTFDPSAYSWQDDAWQGRELQGAVIYELHLGTFTPEGT 131

Qy    119 FEGVIRKLDYLKDLGITAIEIMPIAQFPGKRDWGYDGVYLYAVQNSYGGPEGFRKLVDEA 178
           |    ||||| ||:  ||::|:  | |  :|||||  :||  :||||| ::: || |
Db    132 LEAAAGKLDYLAGLGVDFIELLPVNAFNGTHNWGYDGVQWFAVHEAYGGPEAYQRFVDAA 191

Qy    179 HKKGLGVILDVVYNHVGPEGNYMVKLGPYFSQKYKTPWGLTFNFDDAESDEVRKFILENV 238
          |  ||||| |||||:|| |||: : |||  |      || : | |   || ||::||:|:
Db    192 HAAGLGVIQDVVYNHLGPSGNYLPRFGPYLKQGEGNTWGDSVNLDGPGSDHVRRYILDNL 251

Qy    239 EYWIKEYNVDGFRLDAVHAIIDTSPKHILEEIADVVHKYNRIV------IAESDLNDPRV 292
          |:::| ||| |||||||: |    ||||:   :  : :  |      |||||||:||:
Db    252 AMWLRDYRVDGLRLDAVHALKDERAVHILEDFGALADQISAEVGRPLTLIAESDLNNPRL 311

Qy    293 VNPKEKCGYNIDAQWVDDFHHSIHAYLTGERQGYYTDFGNLDDIVKSYKDVFVYDGKYSN 352
          : |::  || :: || |||||::|  :||| |||:|| :|  : |  :| | :|| ||:
Db    312 LYPRDVNGYGLEGQWSDDFHHAVHVNVTGETTGYYSDFDSLAALAKVLRDGFFHDGSYSS 371

Qy    353 FRRKTHGEPVGELDGCNF--------VVYIQNHDQVGNRGKGERIIKLVDRESYKIAAAL 404
          || : || |:      ||         ||  |||||:|||  |:|: : :   |  :|| |
Db    372 FRERHHGRPI------NFSAVHPAALVVCSQNHDQIGNRATGDRLSQTLPYGSLALAAVL 425

Qy    405 YLLSPYIPMIFMGEEYGEENPFYFFSDFSDSKLIQGVREGRKKE---NGQD----TDPQD 457
          |  |: ||: ||||||   |: ||:   : :| :   ||| ||    | |     ||||
Db    426 TLTGPFTPMLLMGEEYGASTPWQFFTSHPEPELGKATAEGRIKEFERMGWDPAVVPDPQD 485

Qy    458 ESTFNASKLSWKIDEE-----IFSFYKILIKMRKELSIAC-----DRRVNVVNGENWLII 507
          || ||| |    |     :   |: |  :|:          | :|       ||
Db    486 PETFRRSKLDWAEAAEGDHARLLELYRSLTALRRSTPDLTKLGFEDTQVAFDEDARWLRF 545

Qy    508 KGREYFSLYVFSKSSIEVKYSGTLLL 533
          :      |  ||:  : :  :|| ||
Db    546 RRGGVQVLLNFSEQPVSLDGAGTALL 571


RESULT  12
US-08-605-501-4
; Sequence 4, Application US/08605501
; Patent No. 5834287
;  GENERAL INFORMATION:
;    APPLICANT:  KUBOTA, Michio
;    APPLICANT:  TSUSAKI, Keiji
;    APPLICANT:  HATTORI, Kazuko
;    APPLICANT:  SUGIMOTO, Toshiyuki
;    TITLE OF INVENTION:  DNA ENCODING ENZYME, RECOMBINANT, DNA AND
;    TITLE OF INVENTION:  ENZYME, TRANSFORMANT, AND THEIR PREPARATIONS AND USE
;    NUMBER OF SEQUENCES:  18
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BROWDY AND NEIMARK
;      STREET:  419 Seventh Street, N.W., Suite 300
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
```

```
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/605,501
;      FILING DATE:  26-FEB-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/399,646
;      FILING DATE:  07-MAR-1995
;      APPLICATION NUMBER:  JP 59834
;      FILING DATE:  07-MAR-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 59840
;      FILING DATE:  07-MAR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BROWDY, Roger L.
;      REGISTRATION NUMBER:  25,618
;      REFERENCE/DOCKET NUMBER:  KUBOTA=5
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-628-5197
;      TELEFAX:  202-737-3528
;      TELEX:  248633
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  597 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-605-501-4


  Query Match             35.6%;  Score 1059;  DB 2;  Length 597;
  Best Local Similarity   40.5%;  Pred. No. 1.4e-83;
  Matches  229;  Conservative   89;  Mismatches  190;  Indels   58;  Gaps   15;

Qy      14 FTLWAPYQKSVKLKVLEKG-LYEMER------DEKGYFTI----TLNNVKVRDRYKYVLD 62
           : :|||  :|| |  |  |  | |:|      :: |::|     |  ||     | |:||
Db      18 YDVWAPNAESVTL--LAGGERYAMQRRAETGPEDAGWWTAAGAPTDGNVD----YGYLLD 71

Qy      63 -DASEIPDPASRYQPEGVHGPSQIIQESK-EFNNETFLKKE--DLIIYEIHVGTFTPEGT 118
            | : :||| :| ||:|||  |:    |   : :: :  :|    :|||:|:|||||||
Db      72 GDETPLPDPRTRRQPDGVHALSRTFDPSAYSWQDDAWQGRELQGAVIYELHLGTFTPEGT 131

Qy     119 FEGVIRKLDYLKDLGITAIEIMPIAQFPGKRDWGYDGVYLYAVQNSYGGPEGFRKLVDEA 178
           |   |||||  ||:  ||::|:  | |  :||||||  :||  :||||| ::: || |
Db     132 LEAAAGKLDYLAGLGVDFIELLPVNAFNGTHNWGYDGVQWFAVHEAYGGPEAYQRFVDAA 191

Qy     179 HKKGLGVILDVVYNHVGPEGNYMVKLGPYFSQKYKTPWGLTFNFDDAESDEVRKFILENV 238
           |  ||||| |||||:|| |||: : |||  |     || : | |   || ||::||:|:
Db     192 HAAGLGVIQDVVYNHLGPSGNYLPRFGPYLKQGEGNTWGDSVNLDGPGSDHVRRYILDNL 251

Qy     239 EYWIKEYNVDGFRLDAVHAIIDTSPKHILEEIADVVHKYNRIV------IAESDLNDPRV 292
             |:::| ||| |||||||: |    ||||:   :  : :  |      |||||||:||:
Db     252 AMWLRDYRVDGLRLDAVHALKDERAVHILEDFGALADQISAEVGRPLTLIAESDLNNPRL 311

Qy     293 VNPKEKCGYNIDAQWVDDFHHSIHAYLTGERQGYYTDFGNLDDIVKSYKDVFVYDGKYSN 352
           : |::  || :: || |||||::|  :|||  |||:|| :|  : |  :| | :|| ||:
Db     312 LYPRDVNGYGLEGQWSDDFHHAVHVNVTGETTGYYSDFDSLAALAKVLRDGFFHDGSYSS 371

Qy     353 FRRKTHGEPVGELDGCNF--------VVYIQNHDQVGNRGKGERIIKLVDRESYKIAAAL 404
           || : || |:       ||         ||  |||||:|||  |:|: : :   |  :|| |
Db     372 FRERHHGRPI------NFSAVHPAALVVCSQNHDQIGNRATGDRLSQTLPYGSLALAAVL 425

Qy     405 YLLSPYIPMIFMGEEYGEENPFYFFSDFSDSKLIQGVREGRKKE---NGQD----TDPQD 457
           |  |: ||: ||||||   |: ||:   : :| :   ||| ||    | |     ||||
Db     426 TLTGPFTPMLLMGEEYGASTPWQFFTSHPEPELGKATAEGRIKEFERMGWDPAVVPDPQD 485

Qy     458 ESTFNASKLSWKIDEE-----IFSFYKILIKMRKELSIAC-----DRRVNVVNGENWLII 507
             ||  ||| |    |      :   |: |  :|:           | :|       ||
Db     486 PETFRRSKLDWAEAAEGDHARLLELYRSLTALRRSTPDLTKLGFEDTQVAFDEDARWLRF 545

Qy     508 KGREYFSLYVFSKSSIEVKYSGTLLL 533
           :      |  ||:  : :  :|| ||
Db     546 RRGGVQVLLNFSEQPVSLDGAGTALL 571


RESULT  13
US-08-399-646-2
; Sequence 2, Application US/08399646
; Patent No. 5556781
;  GENERAL INFORMATION:
;    APPLICANT:  KUBOTA, Michio
;    APPLICANT:  TSUSAKI, Keiji
;    APPLICANT:  HATTORI, Kazuko
;    APPLICANT:  SUGIMOTO, Toshiyuki
;    TITLE OF INVENTION:  DNA ENCODING ENZYME, RECOMBINANT DNA AND
;    TITLE OF INVENTION:  ENZYME, TRANSFORMANT, AND THEIR PREPARATIONS AND USE
;    NUMBER OF SEQUENCES:  18
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BROWDY AND NEIMARK
;      STREET:  419 Seventh Street, N.W., Suite 300
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/399,646
;      FILING DATE:  07-MAR-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 59834
;      FILING DATE:  07-MAR-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 59840
;      FILING DATE:  07-MAR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BROWDY, Roger L.
;      REGISTRATION NUMBER:  25,618
;      REFERENCE/DOCKET NUMBER:  KUBOTA=5
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-628-5197
;      TELEFAX:  202-737-3528
;      TELEX:  248633
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  589 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-399-646-2


  Query Match             34.4%;  Score 1025;  DB 1;  Length 589;
  Best Local Similarity   40.3%;  Pred. No. 1.2e-80;
  Matches  229;  Conservative   80;  Mismatches  197;  Indels   62;  Gaps   16;

Qy      14 FTLWAPYQKSVKLKVLEKG-LYEMER------DEKGYFTITLNNVKVRDRYKYVLD-DAS 65
           | :|||   :| |  |  |  ||| |       ::|::|           | |:|| |
Db      10 FDIWAPEAGTVTL--LAGGERYEMGRRPGNGPADEGWWTAADAPTGADVDYGYLLDGDEI 67

Qy      66 EIPDPASRYQPEGVHGPSQIIQESKEFNNETFLKKEDL---IIYEIHVGTFTPEGTFEGV 122
             :||| :| ||||||  |:         :   :  :|    :|||:|:|||||||| :
Db      68 PLPDPRTRRQPEGVHALSRTFDPGAHRWQDAGWQGRELQGSVIYELHIGTFTPEGTLDAA 127

Qy     123 IRKLDYLKDLGITAIEIMPIAQFPGKRDWGYDGVYLYAVQNSYGGPEGFRKLVDEAHKKG 182
           |||||  |||  ||::|:  | |  :||||||  :||   ||||  ::: || ||  |
Db     128 AGKLDYLAGLGIDFIELLPVNAFNGTHNWGYDGVQWFAVHEGYGGPAAYQRFVDAAHAAG 187

Qy     183 LGVILDVVYNHVGPEGNYMVKLGPYFSQKYKTPWGLTFNFDDAESDEVRKFILENVEYWI 242
           |||| ||||||:|| |||: : |||       || : | |   || ||::||:||  |:
Db     188 LGVIQDVVYNHLGPSGNYLPRYGPYLKHGEGNTWGDSVNLDGPGSDHVRQYILDNVAMWL 247
```

```
Qy    243 KEYNVDGFRLDAVHAIIDTSPKHILEE---IADVVHKYN---RIVIAESDLNDPRVVNPK 296
          ::| ||| |||||||: |    |||||   :|| :         :|||||||:||:: |:
Db    248 RDYRVDGLRLDAVHALKDERAVHILEEFGALADALSSEGGRPLTLIAESDLNNPRLLYPR 307

Qy    297 EKCGYNIDAQWVDDFHHSIHAYLTGERQGYYTDFGNLDDIVKSYKDVFVYDGKYSNFRRK 356
          :  || :  || |||||::|  ::||  |||:|| :|  : |  :| | :|| ||:|| :
Db    308 DVNGYGLAGQWSDDFHHAVHVNVSGETTGYYSDFDSLGALAKVLRDGFFHDGSYSSFRGR 367

Qy    357 THGEPVGELDGCNF--------VVYIQNHDQVGNRGKGERIIKLVDRESYKIAAALYLLS 408
          || |:      ||        ||  |||||:|||  |:|: : :   |  :|| | |
Db    368 CHGRPI------NFSAVHPAALVVCSQNHDQIGNRATGDRLSQSLPYGSLALAAVLTLTG 421

Qy    409 PYIPMIFMGEEYGEENPFYFFSDFSDSKLIQGVREGRKKE---NGQD----TDPQDESTF 461
          |: ||:|||||||   |: ||:   : :| :   ||| :|    | |     ||||  ||
Db    422 PFTPMLFMGEEYGATTPWQFFTSHPEPELGKATAEGRIREFERMGWDPAVVPDPQDPETF 481

Qy    462 NASKLSWK-----IDEEIFSFYKILIKMRK------ELSIACDRRVNVVNGENWLIIKGR 510
            ||| |        :   |: || :|:       |  | |  |   :   ||
Db    482 TRSKLDWAEASAGDHARLLELYRSLITLRRSTPELARLGFA-DTAVEFDDDARWL----- 535

Qy    511 EYFS-----LYVFSKSSIEVKYSGTLLL 533
          |:      :  |:   | :   || ||
Db    536 RYWRGGVQVVLNFADRPISLDRPGTALL 563


RESULT  14
US-08-607-321-2
; Sequence 2, Application US/08607321
; Patent No. 5716813
;  GENERAL INFORMATION:
;    APPLICANT:  KUBOTA, Michio
;    APPLICANT:  TSUSAKI, Keiji
;    APPLICANT:  HATTORI, Kazuko
;    APPLICANT:  SUGIMOTO, Toshiyuki
;    TITLE OF INVENTION:  DNA ENCODING ENZYME, RECOMBINANT DNA AND
;    TITLE OF INVENTION:  ENZYME, TRANSFORMANT, AND THEIR PREPARATIONS AND USE
;    NUMBER OF SEQUENCES:  18
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BROWDY AND NEIMARK
;      STREET:  419 Seventh Street, N.W., Suite 300
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/607,321
;      FILING DATE:  26-FEB-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/399,646
;      FILING DATE:  07-MAR-1995
;      APPLICATION NUMBER:  JP 59834
;      FILING DATE:  07-MAR-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 59840
;      FILING DATE:  07-MAR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BROWDY, Roger L.
;      REGISTRATION NUMBER:  25,618
;      REFERENCE/DOCKET NUMBER:  KUBOTA=5
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-628-5197
;      TELEFAX:  202-737-3528
;      TELEX:  248633
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  589 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-607-321-2


  Query Match             34.4%;  Score 1025;  DB 1;  Length 589;
  Best Local Similarity   40.3%;  Pred. No. 1.2e-80;
  Matches  229;  Conservative   80;  Mismatches  197;  Indels   62;  Gaps   16;

Qy     14 FTLWAPYQKSVKLKVLEKG-LYEMER------DEKGYFTITLNNVKVRDRYKYVLD-DAS 65
          | :|||   :| |  |  |  ||| |       ::|::|           | |:|| |
Db     10 FDIWAPEAGTVTL--LAGGERYEMGRRPGNGPADEGWWTAADAPTGADVDYGYLLDGDEI 67

Qy     66 EIPDPASRYQPEGVHGPSQIIQESKEFNNETFLKKEDL---IIYEIHVGTFTPEGTFEGV 122
           :||| :| ||||||  |:          :   :|   :|||:|:||||||| :
Db     68 PLPDPRTRRQPEGVHALSRTFDPGAHRWQDAGWQGRELQGSVIYELHIGTFTPEGTLDAA 127

Qy    123 IRKLDYLKDLGITAIEIMPIAQFPGKRDWGYDGVYLYAVQNSYGGPEGFRKLVDEAHKKG 182
           |||||  |||  ||:|:  | |  :||||||  :||   ||||  ::: || ||  |
Db    128 AGKLDYLAGLGIDFIELLPVNAFNGTHNWGYDGVQWFAVHEGYGGPAAYQRFVDAAHAAG 187

Qy    183 LGVILDVVYNHVGPEGNYMVKLGPYFSQKYKTPWGLTFNFDDAESDEVRKFILENVEYWI 242
          |||| ||||||:|| |||: : |||         || : | |   || ||::||:||  |:
Db    188 LGVIQDVVYNHLGPSGNYLPRYGPYLKHGEGNTWGDSVNLDGPGSDHVRQYILDNVAMWL 247

Qy    243 KEYNVDGFRLDAVHAIIDTSPKHILEE---IADVVHKYN---RIVIAESDLNDPRVVNPK 296
          ::| ||| |||||||: |    |||||   :|| :         :|||||||:||:: |:
Db    248 RDYRVDGLRLDAVHALKDERAVHILEEFGALADALSSEGGRPLTLIAESDLNNPRLLYPR 307

Qy    297 EKCGYNIDAQWVDDFHHSIHAYLTGERQGYYTDFGNLDDIVKSYKDVFVYDGKYSNFRRK 356
          :  || :  || |||||::|  ::||  |||:|| :|  : |  :| | :|| ||:|| :
Db    308 DVNGYGLAGQWSDDFHHAVHVNVSGETTGYYSDFDSLGALAKVLRDGFFHDGSYSSFRGR 367

Qy    357 THGEPVGELDGCNF--------VVYIQNHDQVGNRGKGERIIKLVDRESYKIAAALYLLS 408
          || |:      ||        ||  |||||:|||  |:|: : :   |  :|| | |
Db    368 CHGRPI------NFSAVHPAALVVCSQNHDQIGNRATGDRLSQSLPYGSLALAAVLTLTG 421

Qy    409 PYIPMIFMGEEYGEENPFYFFSDFSDSKLIQGVREGRKKE---NGQD----TDPQDESTF 461
          |: ||:|||||||   |: ||:   : :| :   ||| :|    | |     ||||  ||
Db    422 PFTPMLFMGEEYGATTPWQFFTSHPEPELGKATAEGRIREFERMGWDPAVVPDPQDPETF 481

Qy    462 NASKLSWK-----IDEEIFSFYKILIKMRK------ELSIACDRRVNVVNGENWLIIKGR 510
            ||| |        :   |: || :|:       |  | |  |   :   ||
Db    482 TRSKLDWAEASAGDHARLLELYRSLITLRRSTPELARLGFA-DTAVEFDDDARWL----- 535

Qy    511 EYFS-----LYVFSKSSIEVKYSGTLLL 533
          |:      :  |:   | :   || ||
Db    536 RYWRGGVQVVLNFADRPISLDRPGTALL 563


RESULT  15
US-08-961-240-2
; Sequence 2, Application US/08961240
; Patent No. 5830715
;  GENERAL INFORMATION:
;    APPLICANT:  KUBOTA, Michio
;    APPLICANT:  TSUSAKI, Keiji
;    APPLICANT:  HATTORI, Kazuko
;    APPLICANT:  SUGIMOTO, Toshiyuki
;    TITLE OF INVENTION:  DNA ENCODING ENZYME, RECOMBINANT DNA AND
;    TITLE OF INVENTION:  ENZYME, TRANSFORMANT, AND THEIR PREPARATIONS AND USE
;    NUMBER OF SEQUENCES:  18
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BROWDY AND NEIMARK
;      STREET:  419 Seventh Street, N.W., Suite 300
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
```

```
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/961,240
;       FILING DATE:  30-OCT-1997
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/399,646
;       FILING DATE:  07-MAR-1995
;       APPLICATION NUMBER:  JP 59834
;       FILING DATE:  07-MAR-1994
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  JP 59840
;       FILING DATE:  07-MAR-1994
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  BROWDY, Roger L.
;       REGISTRATION NUMBER:  25,618
;       REFERENCE/DOCKET NUMBER:  KUBOTA=5
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  202-628-5197
;       TELEFAX:  202-737-3528
;       TELEX:  248633
;   INFORMATION FOR SEQ ID NO:  2:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  589 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-961-240-2


  Query Match             34.4%;  Score 1025;  DB 2;  Length 589;
  Best Local Similarity   40.3%;  Pred. No. 1.2e-80;
  Matches  229;  Conservative   80;  Mismatches  197;  Indels   62;  Gaps   16;

Qy      14 FTLWAPYQKSVKLKVLEKG-LYEMER------DEKGYFTITLNNVKVRDRYKYVLD-DAS 65
           | :|||   :| |  |  |  ||| |       ::|::|           | |:|| |
Db      10 FDIWAPEAGTVTL--LAGGERYEMGRRPGNGPADEGWWTAADAPTGADVDYGYLLDGDEI 67

Qy      66 EIPDPASRYQPEGVHGPSQIIQESKEFNNETFLKKEDL---IIYEIHVGTFTPEGTFEGV 122
            :||| :| ||||||  |:          :   :  :|   :|||:|:||||||| :
Db      68 PLPDPRTRRQPEGVHALSRTFDPGAHRWQDAGWQGRELQGSVIYELHIGTFTPEGTLDAA 127

Qy     123 IRKLDYLKDLGITAIEIMPIAQFPGKRDWGYDGVYLYAVQNSYGGPEGFRKLVDEAHKKG 182
             |||||  |||  ||::|:  | |  :||||||  :||   ||||  ::: || ||  |
Db     128 AGKLDYLAGLGIDFIELLPVNAFNGTHNWGYDGVQWFAVHEGYGGPAAYQRFVDAAHAAG 187

Qy     183 LGVILDVVYNHVGPEGNYMVKLGPYFSQKYKTPWGLTFNFDDAESDEVRKFILENVEYWI 242
           |||| ||||||:|| |||: : |||         || : | |   || ||::||:||  |:
Db     188 LGVIQDVVYNHLGPSGNYLPRYGPYLKHGEGNTWGDSVNLDGPGSDHVRQYILDNVAMWL 247

Qy     243 KEYNVDGFRLDAVHAIIDTSPKHILEE---IADVVHKYN---RIVIAESDLNDPRVVNPK 296
           ::| ||| |||||||: |    |||||   :|| :          :|||||||:||:: |:
Db     248 RDYRVDGLRLDAVHALKDERAVHILEEFGALADALSSEGGRPLTLIAESDLNNPRLLYPR 307

Qy     297 EKCGYNIDAQWVDDFHHSIHAYLTGERQGYYTDFGNLDDIVKSYKDVFVYDGKYSNFRRK 356
           :  || :  || ||||||::|  ::||  |||:|| :|  : |  :| | :|| ||:|| :
Db     308 DVNGYGLAGQWSDDFHHAVHVNVSGETTGYYSDFDSLGALAKVLRDGFFHDGSYSSFRGR 367

Qy     357 THGEPVGELDGCNF--------VVYIQNHDQVGNRGKGERIIKLVDRESYKIAAALYLLS 408
            || |:      ||        ||  |||||:|||  |:|: : :   |  :|| | |
Db     368 CHGRPI------NFSAVHPAALVVCSQNHDQIGNRATGDRLSQSLPYGSLALAAVLTLTG 421

Qy     409 PYIPMIFMGEEYGEENPFYFFSDFSDSKLIQGVREGRKKE---NGQD----TDPQDESTF 461
           |: ||:|||||||   |: ||:   : :| :   ||| :|    | |     ||||  ||
Db     422 PFTPMLFMGEEYGATTPWQFFTSHPEPELGKATAEGRIREFERMGWDPAVVPDPQDPETF 481

Qy     462 NASKLSWK-----IDEEIFSFYKILIKMRK------ELSIACDRRVNVVNGENWLIIKGR 510
             ||| |          :   |: || :|:       |  | | |   :   ||
Db     482 TRSKLDWAEASAGDHARLLELYRSLITLRRSTPELARLGFA-DTAVEFDDDARWL----- 535

Qy     511 EYFS-----LYVFSKSSIEVKYSGTLLL 533
           |:      :  |:   | :   || ||
```

```
Db     536 RYWRGGVQVVLNFADRPISLDRPGTALL 563


Search completed: July 15, 2002, 12:18:04
Job time: 96 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: July 15, 2002, 12:16:28 ; Search time 32.43 Seconds
(without alignments)
1656.304 Million cell updates/sec

Title: US-09-298-924-6
Perfect score: 2978
Sequence: 1 MTFAYKIDGNEVIFTLWAPY.....PQHIEEGKYEFDKGFALYKL 559

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2973	99.8	558	2	JC5135	alpha-amylase (EC
2	2467	82.8	561	2	S73087	alpha-amylase (EC
3	1748.5	58.7	556	2	JC5132	alpha-amylase (EC
4	1288.5	43.3	620	2	AH1827	alpha-amylase [imp
5	1060.5	35.6	598	2	S65770	maltooligosyltreha
6	1025	34.4	596	2	JC4697	trehalose trehaloh
7	1024.5	34.4	575	2	JC7727	malto-oligosyltreh
8	997.5	33.5	580	2	G70763	probable maltoolig
9	978.5	32.9	594	2	AI0673	probable hydrolase
10	905.5	30.4	600	2	H75516	probable maltoolig
11	881.5	29.6	583	2	G83375	probable glycosyl
12	876.5	29.4	601	2	D96001	probable 1,4-alpha
13	860.5	28.9	595	2	AG3194	alpha-amylase Atu5
14	524.5	17.6	552	2	S19134	probable alpha-glu
15	524.5	17.6	552	2	AH1915	hypothetical prote
16	391	13.1	666	2	B56639	1,4-alpha-glucan b
17	374.5	12.6	627	1	S40048	1,4-alpha-glucan b
18	374.5	12.6	638	2	S18599	1,4-alpha-glucan b
19	365	12.3	1280	2	E95031	alkaline amylopull
20	364	12.2	1256	2	G97902	alpha-amylase (EC
21	361	12.1	630	2	D70363	1,4-alpha-glucan b
22	349.5	11.7	783	2	B84823	probable isoamylas
23	348.5	11.7	666	2	E71565	probable glycogen
24	347.5	11.7	764	2	AG1895	1,4-alpha-glucan b
25	345.5	11.6	422	2	S31839	1,4-alpha-glucan b
26	345.5	11.6	664	2	B86539	glycogen hydrolase
27	345.5	11.6	664	2	B72084	glycosyl hydrolase
28	345.5	11.6	666	2	G81717	glycosyl hydrolase
29	345.5	11.6	707	2	S77094	glycogen operon pr
30	344.5	11.6	741	1	S34218	1,4-alpha-glucan b
31	343	11.5	718	2	S73088	glycogen operon pr
32	340	11.4	730	1	I64118	1,4-alpha-glucan b
33	339.5	11.4	789	2	T01321	probable isoamylas
34	336	11.3	702	2	B85091	isoamylase-like pr
35	335	11.2	764	2	S47569	1,4-alpha-glucan b
36	335	11.2	1072	2	A84112	alkaline amylopull
37	334.5	11.2	717	2	H84057	pullulanase amyX [
38	333.5	11.2	720	2	B75540	glycogen operon pr
39	333.5	11.2	770	2	S76095	hypothetical prote
40	330.5	11.1	562	2	C82263	trehalose-6-phosph
41	329.5	11.1	818	2	T02231	probable isoamylas
42	329	11.0	733	2	T49679	probable branching
43	327.5	11.0	735	2	F98228	1,4-alpha-glucan b
44	327.5	11.0	735	2	AH3057	glycogen branching
45	326.5	11.0	603	2	AC0674	probable glycogen

ALIGNMENTS

RESULT 1
JC5135
alpha-amylase (EC 3.2.1.1) - Sulfolobus solfataricus
C;Species: Sulfolobus solfataricus
C;Date: 31-Jan-1997 #sequence_revision 27-Feb-1997 #text_change 20-Jun-2000
C;Accession: JC5135
R;Kobayashi, K.; Kato, M.; Miura, Y.; Kettoku, M.; Komeda, T.; Iwamatsu, A.
Biosci. Biotechnol. Biochem. 60, 1882-1885, 1996
A;Title: Gene cloning and expression of new trehalose-producing enzymes from the hype
A;Reference number: JC5134; MUID:97141610
A;Accession: JC5135
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-558 <KOB>
A;Cross-references: DDBJ:D64130; NID:g987702; PIDN:BAA11010.1; PID:g1838936
A;Experimental source: strain KM1
C;Comment: This enzyme hydrolyzes starch alpha-1,4 linkage adjacent to the alpha-1,1
C;Superfamily: trehalose trehalohydrolase
C;Keywords: glycosidase; hydrolase

Query Match 99.8%; Score 2973; DB 2; Length 558;
Best Local Similarity 100.0%; Pred. No. 1.1e-185;
Matches 558; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       2 TFAYKIDGNEVIFTLWAPYQKSVKLKVLEKGLYEMERDEKGYFTITLNNVKVRDRYKYVL 61
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 TFAYKIDGNEVIFTLWAPYQKSVKLKVLEKGLYEMERDEKGYFTITLNNVKVRDRYKYVL 60

Qy      62 DDASEIPDPASRYQPEGVHGPSQIIQESKEFNNETFLKKEDLIIYEIHVGTFTPEGTFEG 121
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 DDASEIPDPASRYQPEGVHGPSQIIQESKEFNNETFLKKEDLIIYEIHVGTFTPEGTFEG 120

Qy     122 VIRKLDYLKDLGITAIEIMPIAQFPGKRDWGYDGVYLYAVQNSYGGPEGFRKLVDEAHKK 181
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 VIRKLDYLKDLGITAIEIMPIAQFPGKRDWGYDGVYLYAVQNSYGGPEGFRKLVDEAHKK 180

Qy     182 GLGVILDVVYNHVGPEGNYMVKLGPYFSQKYKTPWGLTFNFDDAESDEVRKFILENVEYW 241
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 GLGVILDVVYNHVGPEGNYMVKLGPYFSQKYKTPWGLTFNFDDAESDEVRKFILENVEYW 240

Qy     242 IKEYNVDGFRLDAVHAIIDTSPKHILEEIADVVHKYNRIVIAESDLNDPRVVNPKEKCGY 301
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 IKEYNVDGFRLDAVHAIIDTSPKHILEEIADVVHKYNRIVIAESDLNDPRVVNPKEKCGY 300

Qy     302 NIDAQWVDDFHHSIHAYLTGERQGYYTDFGNLDDIVKSYKDVFVYDGKYSNFRRKTHGEP 361
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 NIDAQWVDDFHHSIHAYLTGERQGYYTDFGNLDDIVKSYKDVFVYDGKYSNFRRKTHGEP 360

Qy     362 VGELDGCNFVVYIQNHDQVGNRGKGERIIKLVDRESYKIAAALYLLSPYIPMIFMGEEYG 421
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 VGELDGCNFVVYIQNHDQVGNRGKGERIIKLVDRESYKIAAALYLLSPYIPMIFMGEEYG 420
```

```
Qy    422 EENPFYFFSDFSDSKLIQGVREGRKKENGQDTDPQDESTFNASKLSWKIDEEIFSFYKIL 481
Db    421 EENPFYFFSDFSDSKLIQGVREGRKKENGQDTDPQDESTFNASKLSWKIDEEIFSFYKIL 480

Qy    482 IKMRKELSIACDRRVNVVNGENWLIIKGREYFSLYVFSKSSIEVKYSGTLLLSSNNSFPQ 541
Db    481 IKMRKELSIACDRRVNVVNGENWLIIKGREYFSLYVFSKSSIEVKYSGTLLLSSNNSFPQ 540

Qy    542 HIEEGKYEFDKGFALYKL 559
Db    541 HIEEGKYEFDKGFALYKL 558
```

RESULT 2
S73087
alpha-amylase (EC 3.2.1.1) precursor - Sulfolobus solfataricus
N;Alternate names: protein c0620
C;Species: Sulfolobus solfataricus
C;Date: 09-Oct-1997 #sequence_revision 24-Oct-1997 #text_change 20-Jun-2000
C;Accession: S73087
R;Sensen, C.W.; Klenk, H.P.; Singh, R.K.; Allard, G.; Chan, C.C.Y.; Liu, Q.Y.; Penny, S.
Mol. Microbiol. 22, 175-191, 1996
A;Title: Organizational characteristics and information content of an archaeal genome: 1
A;Reference number: S73076; MUID:97055432
A;Accession: S73087
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-561 <SEN>
A;Cross-references: EMBL:Y08256; NID:g1707679; PIDN:CAA69503.1; PID:g1707699
A;Experimental source: strain P2
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1996
C;Superfamily: trehalose trehalohydrolase
C;Keywords: glycosidase; hydrolase

Query Match 82.8%; Score 2467; DB 2; Length 561;
Best Local Similarity 79.7%; Pred. No. 8.6e-153;
Matches 447; Conservative 63; Mismatches 49; Indels 2; Gaps 1;

```
Qy      1 MTFAYKIDGNEVIFTLWAPYQKSVKLKVLEKGLYEMERDEKGYFTITLNNVKVRDRYKYV 60
Db      1 MTFGYKLDEDGVTFNLWAPYQRKVKLKILNRGIYEMERDDKGYFTITLDNVRVGDRYKYI 60

Qy     61 LDDASEIPDPASRYQPEGVHGPSQIIQESKEFNNETF--LKKEDLIIYEIHVGTFTPEGT 118
Db     61 LDDNSEVPDPASRYQPEGVHGYSEIISPDFEWDDENSVKVKREDLVIYELHIGTFTSEGT 120

Qy    119 FEGVIRKLDYLKDLGITAIEIMPIAQFPGKRDWGYDGVYLYAVQNSYGGPEGFRKLVDEA 178
Db    121 FEGVIKKLNYLKELGVTAIEIMPIAQFPGKKDWGYDGVYLYAVQNSYGGPSGFRKLVNEA 180

Qy    179 HKKGLGVILDVVYNHVGPEGNYMVKLGPYFSQKYKTPWGLTFNFDDAESDEVRKFILENV 238
Db    181 HKLGLAVILDVVYNHVGPEGNYMVKLGPYFSEKYKTPWGLTFNFDDAGSDEVRKFILENV 240

Qy    239 EYWIKEYNVDGFRLDAVHAIIDTSPKHILEEIADVVHKYNRIVIAESDLNDPRVVNPKEK 298
Db    241 EYWINEFHVDGFRLDAVHAIIDNSPKHILEDIADVVHKYDKIVIAESDLNDPRVVNPKEK 300

Qy    299 CGYNIDAQWVDDFHHSIHAYLTGERQGYYTDFGNLDDIVKSYKDVFVYDGKYSNFRRKTH 358
Db    301 CGYNIDAQWVDDFHHAIHAFLTGERQGYYSDFGSIGDIVKSYKDVFIYDGKYSNFRRKTH 360

Qy    359 GEPVGELDGCNFVVYIQNHDQVGNRGKGERIIKLVDRESYKIAAALYLLSPYIPMIFMGE 418
Db    361 GKSVGDLDGCKFVVYIQNHDQVGNRGGGERLIKLVDKESYKIAAALYILSPYIPMIFMGE 420

Qy    419 EYGEENPFYFFSDFSDSKLIQGVREGRKKENGQDTDPQDESTFNASKLSWKIDEEIFSFY 478
Db    421 EYGEENPFYYFSDFSDPKLIQGVREGRRRENGQETDPQSDCTFNDSKLSWKINDDILSFY 480

Qy    479 KILIKMRKELSIACDRRVNVVNGENWLIIKGREYFSLYVFSKSSIEVKYSGTLLLSSNNS 538
Db    481 KSLIKIRKEYGLACNRKLSVENGNYWLTVKGNGCLAVYVFSKSVIEMKYSGTLVLSSNSS 540

Qy    539 FPQHIEEGKYEFDKGFALYKL 559
Db    541 FPSQITESKYELDKGFALYKL 561
```

RESULT 3
JC5132
alpha-amylase (EC 3.2.1.1) - Sulfolobus acidocaldarius
N;Alternate names: glycogenase
C;Species: Sulfolobus acidocaldarius
C;Date: 31-Jan-1997 #sequence_revision 27-Feb-1997 #text_change 20-Jun-2000
C;Accession: JC5132
R;Kobayashi, K.; Kato, M.; Miura, Y.; Kettoku, M.; Komeda, T.; Iwamatsu, A.
Biosci. Biotechnol. Biochem. 60, 1720-1723, 1996
A;Title: Gene analysis of trehalose-producing enzymes from hyperthermophilic archaea
A;Reference number: JC5131; MUID:97141330
A;Accession: JC5132
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-556 <KOB>
A;Cross-references: DDBJ:D64131; NID:g987704; PIDN:BAA11011.1; PID:g987705
A;Experimental source: strain ATCC33909
C;Comment: This enzyme hydrolyzes glycosyltrehalose to liberate trehalose.
C;Superfamily: trehalose trehalohydrolase
C;Keywords: glycosidase; hydrolase

Query Match 58.7%; Score 1748.5; DB 2; Length 556;
Best Local Similarity 58.9%; Pred. No. 4.5e-106;
Matches 329; Conservative 87; Mismatches 136; Indels 7; Gaps 5;

```
Qy      2 TFAYKIDGNEVIFTLWAPYQKSVKLKVLEKGLYEMERDEKGYFTITLNNVKVRDRYKYVL 61
Db      3 SFGGNIEKNKGIFKLWAPYVNSVKLK-LSKKLIPMEKNDEGFFEVEIDDIEENLTYSYII 61

Qy     62 DDASEIPDPASRYQPEGVHGPSQIIQESKEFNNETFLKKEDLIIYEIHVGTFTPEGTFEG 121
Db     62 EDKREIPDPASRYQPLGVHDKSQLIRTDYQILDLGKVKIEDLIIYELHVGTFSQEGNFKG 121

Qy    122 VIRKLDYLKDLGITAIEIMPIAQFPGKRDWGYDGVYLYAVQNSYGGPEGFRKLVDEAHKK 181
Db    122 VIEKLDYLKDLGITGIELMPVAQFPGNRDWGYDGVFLYAVQNTYGGPWELAKLVNEAHKR 181

Qy    182 GLGVILDVVYNHVGPEGNYMVKLGPYFSQKYKTPWGLTFNFDDAESDEVRKFILENVEYW 241
Db    182 GIAVILDVVYNHIGPEGNYLLGLGPYFSDRYKTPWGLTFNFDDRGCDQVRKFILENVEYW 241

Qy    242 IKEYNVDGFRLDAVHAIIDTSPKHILEEIADVVHKYNRIVIAESDLNDPRVVNPKEKCGY 301
Db    242 FKTFKIDGLRLDAVHAIFDNSPKHILQEIAEKAHQLGKFVIAESDLNDPKIV--KDDCGY 299

Qy    302 NIDAQWVDDFHHSIHAYLTGERQGYYTDFGNLDDIVKSYKDVFVYDGKYSNFRRKTHGEP 361
Db    300 KIDAQWVDDFHHAVHAFITKEKDYYYQDFGRIEDIEKTFKDVFVYDGKYSRYRGRTHGAP 359

Qy    362 VGELDGCNFVVYIQNHDQVGNRGKGERIIKLVDRESYKIAAALYLLSPYIPMIFMGEEYG 421
Db    360 VGDLPPRKFVVFIQNHDQVGNRGNGERLSILTDKTTYLMAATLYILSPYIPLIFMGEEYY 419

Qy    422 EENPFYFFSDFSDSKLIQGVREGRKKENGQDTDPQDESTFNASKLSWKIDEEIFSFYKIL 481
Db    420 ETNPFFFFSDFSDPVLIKGVREGRLKENNQMIDPQSEEAFLKSKLSWKIDEEVLDYYKQL 479

Qy    482 IKMRKELSIACDRRVNVVNGENWLIIKGREYFSLYVFSKSSIEVKYSGTLLLSSNNSFPQ 541
Db    480 INIRKRYN-NCKRVKEVRREGNCITLIMEKIGIIASFDDIVINSKITGNLLIGI--GFPK 536

Qy    542 HIEEGK-YEFDKGFALYKL 559
Db    537 KLKKDELIKVNRGVGVYQL 555
```

RESULT 4
AH1827
alpha-amylase [imported] - Anabaena sp. (strain PCC 7120)
C;Species: Anabaena sp.
A;Note: Anabaena sp. (strain PCC 7120) is a synonym of Nostoc sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 11-Jan-2002
C;Accession: AH1827
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AH1827
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-620 <KUR>
A;Cross-references: GB:BA000019; PIDN:BAB77692.1; PID:g17135146; GSPDB:GN00179
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: all0168
C;Superfamily: trehalose trehalohydrolase

Query Match 43.3%; Score 1288.5; DB 2; Length 620;
Best Local Similarity 44.0%; Pred. No. 4.2e-76;
Matches 255; Conservative 100; Mismatches 178; Indels 47; Gaps 10;

```
Qy       4 AYKIDGNEVIFTLWAPYQKSVKLKVL--EKGLYEMERDEKGYFTITLNNVKVRDRYKYVL 61
           |: :   |  ||:|:|   || :::|  |: |  ::   :||: | :|:|     |:| |
Db      19 AHYLGNGECEFTVWSPTLNSVAVQILKPEEKLIPLKAQGEGYWQIKVNDVYPGTLYRYQL 78

Qy      62 DDASEIPDPASRYQPEGVHGPSQIIQESKEFNNETF--LKKEDLIIYEIHVGTFTPEGTF 119
           :|     ||||:|||||||| ||::    |: ::|:  :  | :| ||:|||||||||
Db      79 NDQEAFADPASQYQPEGVHGASQVVDHKFEWTDKTWSGISLESMIFYELHVGTFTPEGTF 138

Qy     120 EGVIRKLDYLKDLGITAIEIMPIAQFPGK---------RDWGYDGVYLYAVQNSYGGPEG 170
             :| :|  |::||| |||:||||||||          |:|||||| :|||||| |
Db     139 TTIIPRLPELRELGINAIELMPIAQFPGDDHIEPDLAYRNWGYDGVYPFAVQNSYGSPAD 198

Qy     171 FRKLVDEAHKKGLGVILDVVYNHVGPEGNYMVKLGPYFSQKYKTPWGLTFNFDDAESDEV 230
             :  |:  |: |: |:|||||| |||||| :  |||:: |||||   |||||  |  |
Db     199 LKNFVNACHENGIAVVLDVVYNHFGPEGNYMGQFAPYFTRTYKTPWGNAMNFDDAYSQGV 258

Qy     231 RKFILENVEYWIKEYNVDGFRLDAVHAIIDTSPKHILEEIADVVHKYN------RIVIAE 284
           | : ::|  ||: |:::|| |||| || |   || | |:|: || ::      | :|||
Db     259 RNYFIQNALYWLGEFHIDGLRLDAIQAIYDLGAKHFLWELAEAVHNFSQGGTWKRHLIAE 318

Qy     285 SDLNDPRVVNPKEKCGYNIDAQWVDDFHHSIHAYLTGERQGYYTDFGNLDDIVKSYKDVF 344
           ||||:|::: | |  || :|||| |||||::|| |||:|||||  |||   |: |:| | |
Db     319 SDLNNPQIIRPVESGGYGLDAQWSDDFHHALHALLTGDRQGYYQDFGKCADLAKAYADTF 378

Qy     345 VYDGKYSNFRRKTHGEPVGELDGCNFVVYIQNHDQVGNRGKGERIIKLVDRESYKIAAAL 404
           ||| :|:  |:: ||    :     | | |||||:||: :|||: : :     |:||
Db     379 VYDWRYAPHRKRFHGISCRDRPLSQFSVCIQNHDQIGNQMQGERLSERISFAGLKLAAGA 438

Qy     405 YLLSPYIPMIFMGEEYGEENPFYFFSDFSDSKLIQGVREGRKKE-----NGQD-TDPQDE 458
           |||||:|::|||||||  || :|   ||  ||| || |||:|       :|  ||:
Db     439 VLLSPYLPLLFMGEEYGETAPFIYFVSHSDLDLIQAVRAGRKEEFEAFHYAEDPPDPESA 498

Qy     459 STFNASKLSWKIDEE-----IFSFYKILIKMRK-----------ELSIACDRRVNVVNGE 502
           ||   ||:|::  :     :: :|: || :||            :    |    :|
Db     499 ETFLRCKLNWELRHQGQHKVLWDWYRQLIHLRKTHPALLNYDRDNIEATSDEDKQIVVVR 558

Qy     503 NWLIIKGREYFSLYVFSKSSIE----VKYSGTLLLSSNNS 538
           |   : ||      |: | :     :: |   || | :|
Db     559 RW--CESREVILAMNFNTSPVSLVLTIEKSARKLLDSADS 596
```

RESULT 5
S65770
maltooligosyltrehalose trehalohydrolase - Arthrobacter sp. (strain Q36)
C;Species: Arthrobacter sp.
A;Variety: strain Q36
C;Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 20-Jun-2000
C;Accession: S65770
R;Maruta, K.; Hattori, K.; Nakada, T.; Kubota, M.; Sugimoto, T.; Kurimoto, M.
Biochim. Biophys. Acta 1289, 10-13, 1996
A;Title: Cloning and sequencing of trehalose biosynthesis genes from Arthrobacter sp.
A;Reference number: S65769; MUID:96195835
A;Accession: S65770
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-598 <MAR>
A;Cross-references: EMBL:D63343; NID:g1255444; PIDN:BAA09668.1; PID:g1255446
C;Superfamily: trehalose trehalohydrolase

Query Match 35.6%; Score 1060.5; DB 2; Length 598;
Best Local Similarity 39.7%; Pred. No. 2.7e-61;
Matches 232; Conservative 92; Mismatches 197; Indels 63; Gaps 16;

```
Qy       1 MTFAYKIDGNEVI-----FTLWAPYQKSVKLKVLEKG-LYEMER------DEKGYFTI-- 46
           ||  |  :  : :      : :|||  :|| |  |  |  | |:|       :: |::|
Db       1 MTHTYPREAAKPVLGPARYDVWAPNAESVTL--LAGGERYAMQRRAETGPEDAGWWTAAG 58

Qy      47 --TLNNVKVRDRYKYVLD-DASEIPDPASRYQPEGVHGPSQIIQESK-EFNNETFLKKE- 101
             |  ||     | |:|| | : :||| :| ||:|||  |:    |   : :: :  :|
Db      59 APTDGNVD----YGYLLDGDETPLPDPRTRRQPDGVHALSRTFDPSAYSWQDDAWQGREL 114

Qy     102 -DLIIYEIHVGTFTPEGTFEGVIRKLDYLKDLGITAIEIMPIAQFPGKRDWGYDGVYLYA 160
              :|||:|:|||||||| |    |||||  ||:  ||::|:  | |  :||||||  :|
Db     115 QGAVIYELHLGTFTPEGTLEAAAGKLDYLAGLGVDFIELLPVNAFNGTHNWGYDGVQWFA 174

Qy     161 VQNSYGGPEGFRKLVDEAHKKGLGVILDVVYNHVGPEGNYMVKLGPYFSQKYKTPWGLTF 220
           |  :||||| ::: || ||  ||||| ||||||:|| |||: : |||  |     || :
Db     175 VHEAYGGPEAYQRFVDAAHAAGLGVIQDVVYNHLGPSGNYLPRFGPYLKQGEGNTWGDSV 234

Qy     221 NFDDAESDEVRKFILENVEYWIKEYNVDGFRLDAVHAIIDTSPKHILEEIADVVHKYNRI 280
           | |   || ||::||:|:  |:::| ||| |||||||: |    ||||:   :  : :
Db     235 NLDGPGSDHVRRYILDNLAMWLRDYRVDGLRLDAVHALKDERAVHILEDFGALADQISAE 294

Qy     281 V------IAESDLNDPRVVNPKEKCGYNIDAQWVDDFHHSIHAYLTGERQGYYTDFGNLD 334
           |      |||||||:||:: |::  || :: || ||||||::|  :|||  |||:|| :|
Db     295 VGRPLTLIAESDLNNPRLLYPRDVNGYGLEGQWSDDFHHAVHVNVTGETTGYYSDFDSLA 354

Qy     335 DIVKSYKDVFVYDGKYSNFRRKTHGEPVGELDGCNF--------VVYIQNHDQVGNRGKG 386
           : |  :| | :|| ||:|| : || |:      ||        ||  |||||:|||  |
Db     355 ALAKVLRDGFFHDGSYSSFRERHHGRPI------NFSAVHPAALVVCSQNHDQIGNRATG 408

Qy     387 ERIIKLVDRESYKIAAALYLLSPYIPMIFMGEEYGEENPFYFFSDFSDSKLIQGVREGRK 446
           :|: : :   |  :|| | |  |: ||: ||||||   |: ||:   : :| :   |||
Db     409 DRLSQTLPYGSLALAAVLTLTGPFTPMLLMGEEYGASTPWQFFTSHPEPELGKATAEGRI 468

Qy     447 KE---NGQD----TDPQDESTFNASKLSWKIDEE-----IFSFYKILIKMRKELSIAC-- 492
           ||    | |     ||||  ||  ||| |    |     :   |: | :|:
Db     469 KEFERMGWDPAVVPDPQDPETFRRSKLDWAEAAEGDHARLLELYRSLTALRRSTPDLTKL 528

Qy     493 ---DRRVNVVNGENWLIIKGREYFSLYVFSKSSIEVKYSGTLLL 533
              | :|      ||  :      |  ||:  : :  :|| ||
Db     529 GFEDTQVAFDEDARWLRFRRGGVQVLLNFSEQPVSLDGAGTALL 572
```

RESULT 6
JC4697
trehalose trehalohydrolase (EC 2.4.1.-) - Rhizobium sp.
C;Species: Rhizobium sp.
C;Date: 10-May-1996 #sequence_revision 19-Jul-1996 #text_change 20-Jun-2000
C;Accession: JC4697
R;Maruta, K.; Hattori, K.; Nakada, T.; Kubota, M.; Sugimoto, T.; Kurimoto, M.
Biosci. Biotechnol. Biochem. 60, 717-720, 1996
A;Title: Cloning and sequencing of trehalose biosynthesis genes from Rhizobium sp. M-
A;Reference number: JC4696; MUID:96219094
A;Accession: JC4697
A;Molecule type: DNA

A;Residues: 1-596 <MAR>
A;Cross-references: DDBJ:D78001; NID:g1040695; PIDN:BAA11187.1; PID:g1040697
C;Comment: This enzyme belongs to the alpha-amylase family. It hydrolyzes the products o
oligosyl group and trehalose.
C;Genetics:
A;Gene: MTHase
C;Superfamily: trehalose trehalohydrolase
C;Keywords: glycosyltransferase; hexosyltransferase

Query Match 34.4%; Score 1025; DB 2; Length 596;
Best Local Similarity 40.3%; Pred. No. 5.5e-59;
Matches 229; Conservative 80; Mismatches 197; Indels 62; Gaps 16;

```
Qy      14 FTLWAPYQKSVKLKVLEKG-LYEMER------DEKGYFTITLNNVKVRDRYKYVLD-DAS 65
           | :|||   :| |  |  |  ||| |       ::|::|           | :|:|| |
Db      17 FDIWAPEAGTVTL--LAGGERYEMGRRPGNGPADEGWWTAADAPTGADVDYGYLLDGDEI 74

Qy      66 EIPDPASRYQPEGVHGPSQIIQESKEFNNETFLKKEDL---IIYEIHVGTFTPEGTFEGV 122
           :||| :| |||||  |:          :   :  :|   :|||:|:||||||| :
Db      75 PLPDPRTRRQPEGVHALSRTFDPGAHRWQDAGWQGRELQGSVIYELHIGTFTPEGTLDAA 134

Qy     123 IRKLDYLKDLGITAIEIMPIAQFPGKRDWGYDGVYLYAVQNSYGGPEGFRKLVDEAHKKG 182
           |||||  |||  ||::|:  | |  :||||||  :||   ||||  ::: || ||  |
Db     135 AGKLDYLAGLGIDFIELLPVNAFNGTHNWGYDGVQWFAVHEGYGGPAAYQRFVDAAHAAG 194

Qy     183 LGVILDVVYNHVGPEGNYMVKLGPYFSQKYKTPWGLTFNFDDAESDEVRKFILENVEYWI 242
           |||| ||||||:|| |||: : |||        || : | |   || ||::||:||  |:
Db     195 LGVIQDVVYNHLGPSGNYLPRYGPYLKHGEGNTWGDSVNLDGPGSDHVRQYILDNVAMWL 254

Qy     243 KEYNVDGFRLDAVHAIIDTSPKHILEE---IADVVHKYN---RIVIAESDLNDPRVVNPK 296
           ::| ||| |||||||: |    |||||   :|| :         :|||||||:||:: |:
Db     255 RDYRVDGLRLDAVHALKDERAVHILEEFGALADALSSEGGRPLTLIAESDLNNPRLLYPR 314

Qy     297 EKCGYNIDAQWVDDFHHSIHAYLTGERQGYYTDFGNLDDIVKSYKDVFVYDGKYSNFRRK 356
           :  || :  || |||||::|  ::||  |||:|| :|  : |  :| | :|| ||:|| :
Db     315 DVNGYGLAGQWSDDFHHAVHVNVSGETTGYYSDFDSLGALAKVLRDGFFHDGSYSSFRGR 374

Qy     357 THGEPVGELDGCNF--------VVYIQNHDQVGNRGKGERIIKLVDRESYKIAAALYLLS 408
           || |:      ||        ||  |||||:|||  |:|: : :   |  :|| | |
Db     375 CHGRPI------NFSAVHPAALVVCSQNHDQIGNRATGDRLSQSLPYGSLALAAVLTLTG 428

Qy     409 PYIPMIFMGEEYGEENPFYFFSDFDSDSKLIQGVREGRKKE---NGQD----TDPQDESTF 461
           |: ||:|||||||   |: ||:   : :| :   ||| :|    | |     ||||  ||
Db     429 PFTPMLFMGEEYGATTPWQFFTSHPEPELGKATAEGRIREFERMGWDPAVVPDPQDPETF 488

Qy     462 NASKLSWK-----IDEEIFSFYKILIKMRK------ELSIACDRRVNVVNGENWLIIKGR 510
           |||  |          :   |: || :|:       |  | |  |   :   ||
Db     489 TRSKLDWAEASAGDHARLLELYRSLITLRRSTPELARLGFA-DTAVEFDDDARWL----- 542

Qy     511 EYFS-----LYVFSKSSIEVKYSGTLLL 533
           |:      :  |:   | :   || ||
Db     543 RYWRGGVQVVLNFADRPISLDRPGTALL 570
```

RESULT 7
JC7727
malto-oligosyltrehalose trehalohydrolase (EC 3.2.1.141) - Arthrobacter ramosus
C;Species: Arthrobacter ramosus
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 31-Dec-2001
C;Accession: JC7727; PC7175
R;Yamamoto, T.; Maruta, K.; Watanabe, H.; Yamashita, H.; Kubota, M.; Fukuda, S.; Kurimot
Biosci. Biotechnol. Biochem. 65, 1419-1423, 2001
A;Title: Trehalose-producing operon treYZ from Arthrobacter ramosus S34.
A;Reference number: JC7726; MUID:21364352; PMID:11471747
A;Accession: JC7727
A;Molecule type: DNA
A;Residues: 1-575 <YAM>
A;Cross-references: DDBJ:AB045141
A;Experimental source: strain S34
A;Accession: PC7175
A;Molecule type: protein
A;Residues: 1-20;31-50;298-319 <YA2>

C;Comment: This enzyme is involved in producing trehalose from glycogen.
C;Genetics:
A;Gene: treZ
C;Superfamily: trehalose trehalohydrolase
C;Keywords: glycosidase; hydrolase

Query Match 34.4%; Score 1024.5; DB 2; Length 575;
Best Local Similarity 39.4%; Pred. No. 5.6e-59;
Matches 215; Conservative 90; Mismatches 189; Indels 51; Gaps 13;

```
Qy      14 FTLWAPYQKSVKLKVLEKGLYE--MERDEKGYFTITL---NNVKVRDRYKYVLDDASEIP 68
           | :|||    | | |: :|  |  : ||| |:: :         : | | |::|
Db       5 FPVWAPQAAQVTL-VVGQGRAELPLTRDENGWWALQQPWDGGPDLVD-YGYLVDGKGPFA 62

Qy      69 DPASRYQPEGVHGPSQIIQESKEFNNETFLKKED---------LIIYEIHVGTFTPEGTF 119
           || |  || |||      :  :||:   :   :|          :|||:||||||||||
Db      63 DPRSLRQPRGVH------ELGREFDPARYAWGDDGWRGRDLTGAVIYELHVGTFTPEGTL 116

Qy     120 EGVIRKLDYLKDLGITAIEIMPIAQFPGKRDWGYDGVYLYAVQNSYGGPEGFRKLVDEAH 179
           :  ||:||:|  ||: |:|::|:  | |   ||||||  |||   ||||| ::: ||  |
Db     117 DSAIRRLDHLVRLGVDAVELLPVNAFNGTHGWGYDGVLWYAVHEPYGGPEAYQRFVDACH 176

Qy     180 KKGLGVILDVVYNHVGPEGNYMVKLGPYFSQKYKTPWGLTFNFDDAESDEVRKFILENVE 239
           :|| |: ||||||:|| ||::   |||         ||   | |   |||||::|::|
Db     177 ARGLAVVQDVVYNHLGPSGNHLPDFGPYLGSGAANTWGDALNLDGPLSDEVRRYIIDNAV 236

Qy     240 YWIKEYNVDGFRLDAVHAIIDTSPKHILEEIA----DVVHKYNR--IVIAESDLNDPRVV 293
           ||::: : || |||||||: |    |:|||:|    ::  :  |   :|||||||:::
Db     237 YWLRDMHADGLRLDAVHALRDARALHLLEELAARVDELAGELGRPLTLIAESDLNDPKLI 296

Qy     294 NPKEKCGYNIDAQWVDDFHHSIHAYLTGERQGYYTDFGNLDDIVKSYKDVFVYDGKYSNF 353
           :   || :|||| || ||::|| :|||  ||| ||| |  :|| ::  : :|| :|:|
Db     297 RSRAAHGYGLDAQWDDDVHHAVHANVTGETVGYYADFGGLGALVKVFQRGWFHDGTWSSF 356

Qy     354 RRKTHGEPVG-ELDGCNFVVYIQNHDQVGNRGKGERIIKLVDRESYKIAAALYLLSPYIP 412
           | : || |:  ::      | : |:|||||||  |:|:   |   |   |||| || |: |
Db     357 RERHHGRPLDPDIPFRRLVAFAQDHDQVGNRAVGDRMSAQVGEGSLAAAAALVLLGPFTP 416

Qy     413 MIFMGEEYGEENPFYFFSDFDSDSKLIQGVREGRKKE---NGQD----TDPQDESTFNASK 465
           |:||||||:|   |: ||:   : :| :    ||  |    | |     |||| :||  |
Db     417 MLFMGEEWGARTPWQFFTSHPEPELGEATARGRIAEFARMGWDPAVVPDPQDPATFARSH 476

Qy     466 LSWKIDEE-----IFSFYKILIKMRKELSIACDRRVNVVNGENWLIIKGREYFSLYVFSK 520
           | |   |      : :||  || :|:|| :    |           :   |   :: ||:
Db     477 LDWSEPEREPHAGLLAFYTDLIALRRELPVDAPARE----------VDADEARGVFAFSR 526

Qy     521 SSIEV 525
           : |
Db     527 GPLRV 531
```

RESULT 8
G70763
probable maltooligosyltrehalose trehalohydrolase - Mycobacterium tuberculosis (strain
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C;Accession: G70763
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd,
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Sqares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno
A;Reference number: A70500; MUID:98295987
A;Accession: G70763
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-580 <COL>
A;Cross-references: GB:Z74020; GB:AL123456; NID:g3261584; PIDN:CAA98329.1; PID:g14034
A;Experimental source: strain H37Rv
C;Genetics:
A;Gene: glgZ

```
C;Superfamily: trehalose trehalohydrolase

  Query Match             33.5%;  Score 997.5;  DB 2;  Length 580;
  Best Local Similarity   38.3%;  Pred. No. 3.2e-57;
  Matches  212;  Conservative   91;  Mismatches  215;  Indels   35;  Gaps   11;

Qy      14 FTLWAPYQKSVKLKVLEKGLYEMERDEKGYFTITLNNVKVRDRYKYVL-DDASEIPDPAS 72
           | :|||    |:| |    :: | |   |::  |:       || |:| || : :||| |
Db       4 FRVWAPKPALVRLDV-NGAVHAMTRSADGWWHTTV-AAPADARYGYLLDDDPTVLPDPRS 61

Qy      73 RYQPEGVHGPSQIIQESKEF---NNET---FLKKEDLIIYEIHVGTFTPEGTFEGVIRKL 126
             ||:||| ||  :  :|    :|      |  :|||:|:|||| |||:  | ||
Db      62 ARQPDGVHARSQRWEPPGQFGAARTDTGWPGRSVEGAVIYELHIGTFTTAGTFDAAIEKL 121

Qy     127 DYLKDLGITAIEIMPIAQFPGKRDWGYDGVYLYAVQNSYGGPEGFRKLVDEAHKKGLGVI 186
           ||| |||| :|:||:  | | | |||||| |:|   ||||:|  : :|  | : |||:
Db     122 DYLVDLGIDFVELMPVNSFAGTRGWGYDGVLWYSVHEPYGGPDGLVRFIDACHARRLGVL 181

Qy     187 LDVVYNHVGPEGNYMVKLGPYFSQKYKTPWGLTFNFDDAESDEVRKFILENVEYWIKEYN 246
           :| |:||:|| |||: : ||| |     |||   |   |:|||| :|::    |:::::
Db     182 IDAVFNHLGPSGNYLPRFGPYLSSA-SNPWGDGINIAGADSDEVRHYIIDCALRWMRDFH 240

Qy     247 VDGFRLDAVHAIIDTSPKHILEEIADVVHKYN------RIVIAESDLNDPRVVNPKEKCG 300
            || |||||||::||:  |:|||:|:     :        :|||:| ||||::       |
Db     241 ADGLRLDAVHALVDTTAVHVLEELANATRWLSGQLGRPLSLIAETDRNDPRLITRPSHGG 300

Qy     301 YNIDAQWVDDFHHSIHAYLTGERQGYYTDFGNLDDIVKSYKDVFVYDGKYSNFRRKTHGE 360
           | | ||| || ||:||  ::||||||| |||:|  :  : :: : : | ||:|||: ||
Db     301 YGITAQWNDDIHHAIHTAVSGERQGYYADFGSLATLAYTLRNGYFHAGTYSSFRRRRHGR 360

Qy     361 PV--GELDGCNFVVYIQNHDQVGNRGKGERIIKLVDRESYKIAAALYLLSPYIPMIFMGE 418
            :    :     : |   |||||||  |:|  : :       | ||| | |||  |:|||
Db     361 ALDTSAIPATRLLAYTCTHDQVGNRALGDRPSQYLTGGQLAIKAALTLGSPYTAMLFMGE 420

Qy     419 EYGEENPFYFFSDFSDSKLIQGVREGRKKENGQD--------TDPQDESTFNASKLSWK-- 469
           |:|  :|| ||    : :|     |||:|  :          ||||  ||   ||:|
Db     421 EWGASSPFQFFCSHPEPELAHSTVAGRKEEFAEHGWAADDIPDPQDPQTFQRCKLNWAEA 480

Qy     470 ---IDEEIFSFYKILIKMRKELSIACD-----RRVNVVNGENWLIIKGREYFSLYVFSKS 521
                :  ||: || :|   :  |      |:    : |:::  :
Db     481 GSGEHARLHRFYRDLIALRHNEADLADPWLDHLMVDYDEQQRWVVMRRGQLMIACNLGAE 540

Qy     522 SIEVKYSGTLLLS 534
             |  || |:|:
Db     541 PTCVPVSGELVLA 553


RESULT   9
AI0673
probable hydrolase STY1503 [imported] - Salmonella enterica subsp. enterica serovar Typh
C;Species: Salmonella enterica subsp. enterica serovar Typhi
A;Note: this species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 27-Nov-2001
C;Accession: AI0673
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A;Reference number: AB0502; PMID:11677608
A;Accession: AI0673
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-594 <PAR>
A;Cross-references: GB:AL513382; PIDN:CAD01762.1; PID:g16502610; GSPDB:GN00176
C;Genetics:
A;Gene: STY1503
C;Superfamily: trehalose trehalohydrolase

  Query Match             32.9%;  Score 978.5;  DB 2;  Length 594;
  Best Local Similarity   43.3%;  Pred. No. 5.7e-56;
  Matches  209;  Conservative   76;  Mismatches  169;  Indels   29;  Gaps   12;

Qy       7 IDGNEVIFTLWAPYQKSVKLKVLEKGLYEMERDEKGYFTITLNNVKVRDRYKYVLDDASE 66
           |  : | | |||  |: | |::  |   ||:    |:||: :: |     | :|| |
Db      15 IAADVVRFRLWATGQQKVMLRLAGKD-QEMQASGDGWFTLDVSGVTPGTEYNFVLSDGMV 73

Qy      67 IPDPASRYQPEGVHGPSQIIQESKEFNNETFLKK---EDLIIYEIHVGTFTPEGTFEGVI 123
           :|||||| |   |:||| :|          |  |    |  ::||:| |||||||||   |
Db      74 LPDPASRAQKTDVNGPSYVIDPGSYAWRNTGWKGSRWEQAVVYEMHTGTFTPEGTFHTAI 133

Qy     124 RKLDYLKDLGITAIEIMPIAQFPGKRDWGYDGVYLYAVQNSYGGPEGFRKLVDEAHKKGL 183
           || || :||:| ||:||:|| |:| |||||| |||  ::|| |: |:  :| ||  ||
Db     134 AKLPYLAELGVTVIEVMPVAQFGGERGWGYDGVLLYAPHSAYGTPDDFKAFIDTAHGYGL 193

Qy     184 GVILDVVYNHVGPEGNYMVKLGP-YFSQKYKTPWGLTFNFDDAESDEVRKFILENVEYWI 242
           ||||:| || |||||:  | | :| ::  ||||    :|    | ||::|:|   ||:
Db     194 SVILDIVLNHFGPEGNYLPLLAPAFFHKERMTPWGNGIAYD---VDAVRRYIIEAPLYWL 250

Qy     243 KEYNVDGFRLDAVHAIIDTSPKHILEEIADVVHK---YNRIVIAESDLNDPRVVNPKEKC 299
           ||::|| | ||:  | |:| :|:| |||  : :      | :   |  :   ::|:::
Db     251 TEYHLDGLRFDAIDQIEDSSARHVLVEIAQRIREDITDRPIHLTTEDSRNIISLHPRDQD 310

Qy     300 GYN--IDAQWVDDFHHSIHAYLTGERQGYYTDFGNLDD--IVKSYKDVFVYDGKYSNFRR 355
           |      |:| ||||:::| : ||| | || || :  :  : ::  : | | |: |
Db     311 GNAPLFTAEWNDDFHNAVHVFATGETQAYYNDFADAPEKHLARALAEGFAYQGEIS---- 366

Qy     356 KTHGEPVGELD----GCNFVVYIQNHDQVGNRGKGERIIKLVDRESYKIAAALYLLSPYI 411
              ||| |          || :|||||||||| :|:|:| |   |  |:  |  ||||:|
Db     367 PQTGEPRGVKSTGQPPVAFVDFIQNHDQVGNRAQGDRLITLAGAERTKVLLATLLLSPHI 426

Qy     412 PMIFMGEEYGEENPFYFFSDFSDSKLIQGVREGRKKE----NGQDT-DPQDESTFNASKL 466
           |::|||||||  || ||:||    | : |||| ||     |::  ||    ||  |||
Db     427 PLLFMGEEYGESRPFLFFTDF-HGDLARAVREGRAKEFADHAGENVPDPNAPETFQRSKL 485

Qy     467 SWK 469
           :||
Db     486 NWK 488


RESULT  10
H75516
probable maltooligosyltrehalose trehalohydrolase - Deinococcus radiodurans (strain R1
C;Species: Deinococcus radiodurans
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
C;Accession: H75516
R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J
, M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.,
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A;Reference number: A75250; MUID:20036896
A;Accession: H75516
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-600 <WHI>
A;Cross-references: GB:AE001905; GB:AE000513; NID:g6458144; PIDN:AAF10042.1; PID:g645
A;Experimental source: strain R1
C;Genetics:
A;Gene: DR0464
A;Map position: 1
C;Superfamily: trehalose trehalohydrolase

  Query Match             30.4%;  Score 905.5;  DB 2;  Length 600;
  Best Local Similarity   38.7%;  Pred. No. 3.2e-51;
  Matches  205;  Conservative   86;  Mismatches  170;  Indels   69;  Gaps   20;

Qy      14 FTLWAPYQKSVKLKV---------LEKGLYEMERDEKGYFTITLNNVKVRDRYKYVLDDA 64
           | ||    ::| ::|         |  |:||:|             |    || :|||
Db      38 FRLWTSTARTVAVRVNGTEHVMTSLGGGIYELELP-----------VGPGARYLFVLDGV 86

Qy      65 SEIPDPASRYQPEGVHGPSQIIQESKEFNNETF---------LKKEDLIIYEIHVGTFTP 115
```

```
Db     87 -PTPDPYARFLPDGVHGEAEVV----DFG--TFDWTDADWHGIKLADCVFYEVHVGTFTP 139

Qy    116 EGTFEGVIRKLDYLKDLGITAIEIMPIAQFPGKRDWGYDGVYLYAVQNSYGGPEGFRKLV 175

Db    140 EGTYRAAAEKLPYLKELGVTAIQVMPLAAFDGQRGWGYDGAAFYAPYAPYGRPEDLMALV 199

Qy    176 DEAHKKGLGVILDVVYNHVGPEGNYMVKLGP-YFSQKYKTPWGLTFNFDDAESDEVRKFI 234

Db    200 DAAHRLGLGVFLDVVYNHFGPSGNYLSSYAPSYFTDRFSSAWGM--GLDYAE-PHMRRYV 256

Qy    235 LENVEYWIKEYNVDGFRLDAVHAIIDTSPKHILEEIADVVHKY--NRIVIAESDLNDPRV 292

Db    257 TGNARMWLRDYHFDGLRLDATPYMTDDSETHILTELAQEIHELGGTHLLLAEDHRNLPDL 316

Qy    293 VNPKEKCGYNIDAQWVDDFHHSIHAYLTGERQGYYTDF-GNLDDIVKSYKDVFVYDGKYS 351

Db    317 VTVN-----HLDGIWTDDFHHETRVTLTGEQEGYYAGYRGGAEALAYTIRRGWRYEGQFW 371

Qy    352 NFRRKTH--GEPVGELDGCNFVVYIQNHDQVGNRGKGERIIKL--VDRESYKIAAALYLL 407

Db    372 AVKGEEHERGHPSDALEAPNFVYCIQNHDQIGNRPLGERLHQSDGVTLHEYRGAAALLL- 430

Qy    408 SPYIPMIFMGEEYGEENPFYFFSDFSDSKLIQGVREGRKKE-------NGQDT-DPQDES 459

Db    431 -PMTPLLFQGQEWAASTPFQFFSDHA-GELGQAVSEGRKKEFGGFSGFSGEDVPDPQAEQ 488

Qy    460 TFNASKLSWKIDE-----EIFSFYKILIKMRKELSIACDR-RVNVVNGEN 503

Db    489 TFLNSKLNWAEREGGEHARTLRLYRDLLRLRREDPVLHNRQRENLTTGHD 538


RESULT  11
G83375
probable glycosyl hydrolase PA2164 [imported] - Pseudomonas aeruginosa (strain PAO1)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C;Accession: G83375
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A;Reference number: A82950; MUID:20437337
A;Accession: G83375
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-583 <STO>
A;Cross-references: GB:AE004643; GB:AE004091; NID:g9948178; PIDN:AAG05552.1; GSPDB:GN001
A;Experimental source: strain PAO1
C;Genetics:
A;Gene: PA2164
C;Superfamily: trehalose trehalohydrolase


  Query Match             29.6%;  Score 881.5;  DB 2;  Length 583;
  Best Local Similarity   36.3%;  Pred. No. 1.1e-49;
  Matches  203;  Conservative   95;  Mismatches  174;  Indels   87;  Gaps   19;

Qy      3 FAYKIDGN-EVIFTLWAPYQKSVKLKVLEKGLYEMERDEKGYFTITLNNVKVRDRYKYVL 61

Db      7 FGAQFQGNGRTCFGLWAPDAREVRVETADGRDWPLEGSDSGWFEATL-PCPPGTRYRYRI 65

Qy     62 DDASEIPDPASRYQPEGVHGPSQIIQES------KEFNNETFLKKEDLIIYEIHVGTFTP 115

Db     66 DGRPGVPDPASQFQPDGVHGHSQVLDHGTYAWRVDEWRGRPW---HEAVIYELHVGLF-- 120

Qy    116 EGTFEGVIRKLDYLKDLGITAIEIMPIAQFPGKRDWGYDGVYLYAVQNSYGGPEGFRKLV 175

Db    121 -GSYAEVERFLPRLVELGVTAVELMPLGEFPGRRNWGYDGVLPFAPASAYGTPEQLKHLI 179

Qy    176 DEAHKKGLGVILDVVYNHVGPEGNYMVK-LGPYFSQKYKTPWGLTFNFDDAESDEVRKFI 234

Db    180 DSAHGMGLMVFVDVIYNHFGPDGNYLAQYAAAFFRDDRQTPWGQAIDF---RRGEVREFF 236

Qy    235 LENVEYWIKEYNVDGFRLDAVHAIIDTSPKHILEEIA----------DVVHKYNRIVIAE 284

Db    237 YENALMWLLDYRVDGLRFDAVHAIPDSA---FLVEMARRLRGAAGPERHVH-----LVLE 288

Qy    285 SDLNDPRVVNPKEKCGYNIDAQWVDDFHHSIHAYLTGERQGYYTDFGN-LDDIVKSYKDV 343

Db    289 NDDNRASLL----RQGY--DAQWNDDGHHALHVLLTGENDGYYQDYPEPLRCLARCLAEG 342

Qy    344 FVYDGKYSNFRRKTHGEPVGELDGCNFVVYIQNHDQVGNRGKGERIIKLVDRESYKIAAA 403

Db    343 FVYQGE-ANRHGRPRGEPSADLAPDAFVLFLQNHDQVGNRAFGERLSVLAEPQALRLAIA 401

Qy    404 LYLLSPYIPMIFMGEEYGEENPFYFFSDFSDSKLIQGVREGRKKENGQD---------- 452

Db    402 LQLLAPMIPLLFMGEECAAREPFLYFTD-HQGELADAVREGRRKEFGEFGRFGEGATLAS 460

Qy    453 -TDPQDESTFNASK-------LSWKIDEEIFSFYKILIKMRKELSI---------- 490

Db    461 LPDPNAVETFERSRPGLADCDPAWR-----GFYRQLLEIRHEHLIPRLRGARSLGVTTI 514

Qy    491 ---ACDRRVNVVNGENWLI 506

Db    515 AGAALSARWRLGDGSDWRI 533


RESULT  12
D96001
probable 1,4-alpha-glucan branching enzyme (EC 2.4.1.18) [imported] - Sinorhizobium m
C;Species: Sinorhizobium meliloti
C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 14-Sep-2001
C;Accession: D96001
R;Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Her
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A;Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing e
A;Reference number: A95842; MUID:21396508; PMID:11481431
A;Accession: D96001
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-601 <KUR>
A;Cross-references: GB:AL591985; PIDN:CAC49676.1; PID:g15141163; GSPDB:GN00167
A;Experimental source: strain 1021, megaplasmid pSymB
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubl
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.
 L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelau
hebault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh,
A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A;Reference number: A96039; MUID:21368234; PMID:11474104
A;Contents: annotation
C;Genetics:
A;Gene: glgB2; SMb21447
A;Genome: plasmid
C;Superfamily: trehalose trehalohydrolase
C;Keywords: glycosyltransferase; hexosyltransferase


  Query Match             29.4%;  Score 876.5;  DB 2;  Length 601;
  Best Local Similarity   35.3%;  Pred. No. 2.5e-49;
  Matches  207;  Conservative   99;  Mismatches  204;  Indels   77;  Gaps   19;

Qy      7 IDGNEVIFTLWAPYQKSVKLKVLEKGLYEMERDEKGYFTITLNNVKVRDRYKYVLDDASE 66

Db     19 IDSDTCRFRLWAPDEREVDL-VLGGAVHKMQSLDGGWFEITL-AAKAGERYCFRLADGTE 76

Qy     67 IPDPASRYQPEGVHGPSQIIQESK-EFNNETFLKK--EDLIIYEIHVGTFTPEGTFEGVI 123

Db     77 VADPASSAQEREASGTSIVVDHAAYEWQTSSWRGRPWEEAVISELHVGCFTPEGTFRAAI 136

Qy    124 RKLDYLKDLGITAIEIMPIAQFPGKRDWGYDGVYLYAVQNSYGGPEGFRKLVDEAHKKGL 183
```

```
Db    137 ERLPHLAGAGITAIEIMPVAQFPGVRGWGYDGVLHYAPHNAYGKPDDLKALVDAAHSLGL 196

Qy    184 GVILDVVYNHVGPEGNYMVKLGP-YFSQKYKTPWGLTFNFDDAESDEVRKFILENVEYWI 242
Db    197 TVLLDVVYNHFGPEGNYLSRYASRFFNKDRPTPWGASIAF---EEEAVRRYFIENALYWL 253

Qy    243 KEYNVDGFRLDAVHAIIDTSPKHILEEIADVVHK---YNRIVIAESDLNDPRVVNPKEKC 299
Db    254 GHFRFDGLRLDATEQIRDTTKPHFLVALEHEVREAFAERQIHLVLEDAHRRRSLLQRDAS 313

Qy    300 GYNI--DAQWVDDFHHSIHAYLTGERQGYYTDFGN--LDDIVKSYKDVFVYDGKYSNFR- 354
Db    314 GARMLFDAAWNDDLHNALHVVATGETKGHYRLFADEPWGKIRSALAEGFAVPAKEDNFSP 373

Qy    355 --RKTHGEPVGELDGCNFVVYIQNHDQVGNRGKGERIIKLVDRESYKIAAALYLLSPYIP 412
Db    374 EGSRARVPPQGRVN------FLQNHDQIGNRAFGERLASLLQEDSLRVLAAMHMLTPQIP 427

Qy    413 MIFMGEEYGEENPFYFFSDFSDSKLIQGVREGRKKE--------NGQD----TDPQDEST 460
Db    428 LLFMGEEYGETQPFYFFSDY-QGEIAAAIRLGRRDEAENFGGLPEGKTMDDLPDPLDPDV 486

Qy    461 FNASKLSWK-----IDEEIFSFYKILIKMRK---------------------------- 486
Db    487 FAGSKLRWNRATSPAGERHLAYVRDLAVIRQRHIVPLIAGTAVPDGRAYETKDGIIAVDW 546

Qy    487 ELSIAC-DRRVNVVNGENWL-IIKGREYFSLYVFSKSSIEVKYSGTL 531
Db    547 QFGEACLELRVNLLHETHAVPAIRGQPIFT---SETSSVETVYGSEL 590


RESULT  13
AG3194
alpha-amylase Atu5284 [imported] - Agrobacterium tumefaciens (strain C58, Dupont) plasmi
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 01-Feb-2002
C;Accession: AG3194
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L
erage, G.; Gillet, W.; Grant, C.; Guenthner, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; PMID:11743193
A;Accession: AG3194
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-595 <KUR>
A;Cross-references: GB:AE008687; PIDN:AAL45973.1; PID:g17743726; GSPDB:GN00188
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: Atu5284
A;Genome: plasmid
C;Superfamily: trehalose trehalohydrolase


  Query Match             28.9%;  Score 860.5;  DB 2;  Length 595;
  Best Local Similarity   34.7%;  Pred. No. 2.7e-48;
  Matches  206;  Conservative   89;  Mismatches  246;  Indels   53;  Gaps   16;

Qy      4 AYKIDGNEVIFTLWAPYQKSVKLKVLEKGLYEMERDEKGYFTITLNNVKVRDRYKYVLDD 63
Db     12 AQRHQSGDTDFSIWAPASATVKL-WLNDAEFDMHTAGDGWHDIT-KPALPGDRYGFVLAD 69

Qy     64 ASEIPDPASRYQPEGVHGPSQIIQESKEFNNETFLKK--EDLIIYEIHVGTFTPEGTFEG 121
Db     70 RTRVADPASNRQQEGPRGPSLIVNHDFAWKNPNWKGRPWHEAVVYELHIGTFTPEGTFAA 129

Qy    122 VIRKLDYLKDLGITAIEIMPIAQFPGKRDWGYDGVYLYAVQNSYGGPEGFRKLVDEAHKK 181
Db    130 AAEKLEYLADVGITTIELMPLATFAGSRGWGYDGVLQFSPQRDYGTPDELKAFIDQAHGH 189

Qy    182 GLGVILDVVYNHVGPEGNYMVKLGPYFSQKYKTPWGLTFNFDDAESDEVRKFILENVEYW 241
Db    190 GIMVLLDVVYNHFGPAGNTLQAYAPAFFKKHETPWGPAPDFNRA---EVRSFFLQNAFYW 246

Qy    242 IKEYNVDGFRLDAVHAII--DTSPKHILE---EIADVVHKYNRIVIAESDLNDPRVVNPK 296
Db    247 LETYRFDGLRIDAADHLAGGDGEVDFLIEMAREVKRTIRNRHVHLVIEDARNAASPMTPM 306

Qy    297 EKCGYNIDAQWVDDFHHSIHAYLTGERQGYYTDFGN--LDDIVKSYKDVFVYDGKYSNFR 354
Db    307 ADGAILVDAQWNDDFHHVIHVATTNEEGGIYEDFASRPYENLRRSLATGFVYQGEPRPSR 366

Qy    355 R-KTHGEPVGELDGCNFVVYIQNHDQVGNRGKGERIIKLVDRESYKIAAALYLLSPYIPM 413
Db    367 NFAASGEPSGHLPPHRFVNFLHNHDQAGNRLRGERLRALIPPPLFGTLEAILLLCPQTPL 426

Qy    414 IFMGEEYGEENPFYFFSDFSDSKLIQGVREGRKK-ENGQD----------TDPQDESTFN 462
Db    427 VFMGDEHGSANPFFFFSDHPDHNREQEIRNRLKQAESFQGELPPDASQMVMDPNDQHTMQ 486

Qy    463 ASKLSWKIDEEI------FSFYKILIKMR------------KELSIACDRRVNVVNGENW 504
Db    487 LSTLKWTHAETTEGRQSRADMAALLAKRRGHIWPLLCSHFEKGISLDCEPRCLAI---DW 543

Qy    505 LIIKGREYFSLYVFSKSSIEVKYSGTLLLSSNNSFPQHIEEGKYE-FDKGFALY 557
Db    544 HFKAGRLEMRANL-SENMCELPAVKGEILHRNGS----VGNTRYEGYAAQFAIY 592


RESULT  14
S19134
probable alpha-glucanotransferase - Anabaena variabilis
C;Species: Anabaena variabilis
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 17-Mar-2000
C;Accession: S19134
R;Sato, N.
Plant Mol. Biol. 18, 165-170, 1992
A;Title: Cloning of a low-temperature-induced gene lti2 from the cyanobacterium Anaba
A;Reference number: S19133; MUID:92119230
A;Accession: S19134
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-552 <SAT>
A;Cross-references: EMBL:X59279; NID:g39252; PIDN:CAA41970.1; PID:g39253
C;Superfamily: neopullulanase; alpha-amylase core homology


  Query Match             17.6%;  Score 524.5;  DB 2;  Length 552;
  Best Local Similarity   28.0%;  Pred. No. 1.7e-26;
  Matches  148;  Conservative   90;  Mismatches  162;  Indels  129;  Gaps   22;

Qy     14 FTLWAPYQKSVKLKVLEKGLYE--MERDEKGYFTITLNNVKVRD---RYKYVLDDAS--- 65
Db      7 FTLFAPYNKGAALIASFSDWQEIPMKKGDDGYFRTT---VELEDGTYQYKFRVQTRSWFF 63

Qy     66 ------EIPDPASRYQPEGVHGPSQIIQESKEFNN--ETFLKKED---------LIIYEI 108
Db     64 EEDQWVDVTDPYATDIDES-SGKDNSIARIKDGEKIVDTYVWQHDDKPLPADHELVIYEL 122

Qy    109 HVGTFT-------PEGTFEGVIRKLDYLKDLGITAIEIMPIAQFPGKRDWGYDGVYLYAV 161
Db    123 HVGDFSGGEDDPYARGKYKHVIEKLDYLCELGINAIELLPVKEYPGDYSWGYNPRYFFAT 182

Qy    162 QNSYGGPEGFRKLVDEAHKKGLGVILDVVYNHVGPEGNY-MVKLGPYFSQKYKTP---WG 217
Db    183 ESSYGSTADLKKLVDECHQRGIRIIMDGIYNHSEASSPLTQIDHDYWYHHEPRDPDNNWG 242

Qy    218 LTFNF----DDAESDEVRKFILENVEYWIKEYNVDGFRLDAVHAIIDTSPKHIL----EE 269
Db    243 PEFNYEHYDENLETYPARKFIGDTVRYWVGEYHLDGIRYDAARQIANYDFMHWIAQEAKK 302

Qy    270 IADVVHKYNRIVIAESDLNDPRVVN---PKEKCGYNIDAQWVDDFHHSIHAYLTGERQGY 326
```

```
Db    303 TAGAKPFYN---VAEHIPETTSITNLDGPMDGC-------WHDSFYHTIKAHICGDT--- 349

Qy    327 YTDFGNLDDIVKSYKDVFVYDGKYSNFRRKTHGEPVGELDGCNFVVYIQNHD------QV 380
          |  || |::         | |   |           |   | | |: |||      ::
Db    350 -FDLENLKDVI---------DPKRQGF-----------LGATNVVNYLTNHDHDHIMVEL 388

Qy    381 GNRGKGERIIKLVDRESY---KIAAALYLLSPYIPMIFMGEEYGEENPFYFFSDFSDSKL 437
          |||       ::   |::  |:  |: : :  :|:|:||||:||  |
Db    389 GNR-------EIFHDEAFRRAKLGTAILMTAVGVPLIWMGEEFGEYKP------------ 429

Qy    438 IQGVREGRKKENGQDTDPQDESTFNASKLSWKIDEEIFSFYKILIKMRK 486
                        ||:|  : : |   ::  :| ::| || :||
Db    430 ----------------KQQDQSKIDWTLLGNDLNRSLFDYHKGLIGLRK 462


RESULT  15
AH1915
hypothetical protein all0875 [imported] - Anabaena sp. (strain PCC 7120)
C;Species: Anabaena sp.
A;Note: Anabaena sp. (strain PCC 7120) is a synonym of Nostoc sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 11-Jan-2002
C;Accession: AH1915
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
 Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AH1915
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-552 <KUR>
A;Cross-references: GB:BA000019; PIDN:BAB72832.1; PID:g17130220; GSPDB:GN00179
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: all0875
C;Superfamily: neopullulanase; alpha-amylase core homology


  Query Match           17.6%;  Score 524.5;  DB 2;  Length 552;
  Best Local Similarity   28.0%;  Pred. No. 1.7e-26;
  Matches  148;  Conservative   90;  Mismatches  162;  Indels  129;  Gaps   22;

Qy     14 FTLWAPYQKSVKLKVLEKGLYE--MERDEKGYFTITLNNVKVRD---RYKYVLDDAS--- 65
          |||:||| |   |       |  |:: : |||  |   |:: |   :||: :   |
Db      7 FTLFAPYNKGAALIASFSDWQEIPMKKGDDGYFRTT---VELEDGTYQYKFRVQTRSWFF 63

Qy     66 ------EIPDPASRYQPEGVHGPSQIIQESKEFNN--ETFLKKED---------LIIYEI 108
                :: || :    |   |    |   |:     :|:: : |          |:|||:
Db     64 EEDQWVDVTDPYATDIDES-SGKDNSIARIKDGEKIVDTYVWQHDDKPLPADHELVIYEL 122

Qy    109 HVGTFT-------PEGTFEGVIRKLDYLKDLGITAIEIMPIAQFPGKRDWGYDGVYLYAV 161
          |||  |:         | :: || ||||| :||| |||::|: ::||   |||:  | :|
Db    123 HVGDFSGGEDDPYARGKYKHVIEKLDYLCELGINAIELLPVKEYPGDYSWGYNPRYFFAT 182

Qy    162 QNSYGGPEGFRKLVDEAHKKGLGVILDVVYNHVGPEGNY-MVKLGPYFSQKYKTP---WG 217
          ::|||     :||||| |::|: :|:| :|||         :    ::  : : |   ||
Db    183 ESSYGSTADLKKLVDECHQRGIRIIMDGIYNHSEASSPLTQIDHDYWYHHEPRDPDNNWG 242

Qy    218 LTFNF----DDAESDEVRKFILENVEYWIKEYNVDGFRLDAVHAIIDTSPKHIL----EE 269
            ||:    :: |:   |||| : | ||: ||::|| | ||   | :    | :    ::
Db    243 PEFNYEHYDENLETYPARKFIGDTVRYWVGEYHLDGIRYDAARQIANYDFMHWIAQEAKK 302

Qy    270 IADVVHKYNRIVIAESDLNDPRVVN---PKEKCGYNIDAQWVDDFHHSIHAYLTGERQGY 326
           |     ||   :||       : |   | : |       | | |:|:| |:: |:
Db    303 TAGAKPFYN---VAEHIPETTSITNLDGPMDGC-------WHDSFYHTIKAHICGDT--- 349

Qy    327 YTDFGNLDDIVKSYKDVFVYDGKYSNFRRKTHGEPVGELDGCNFVVYIQNHD------QV 380
          |  || |::         | |   |           |   | | |: |||      ::
Db    350 -FDLENLKDVI---------DPKRQGF-----------LGATNVVNYLTNHDHDHIMVEL 388

Qy    381 GNRGKGERIIKLVDRESY---KIAAALYLLSPYIPMIFMGEEYGEENPFYFFSDFSDSKL 437
          |||       ::   |::   |:  |: : :  :|:|:||||:||  |
Db    389 GNR-------EIFHDEAFRRAKLGTAILMTAVGVPLIWMGEEFGEYKP------------ 429

Qy    438 IQGVREGRKKENGQDTDPQDESTFNASKLSWKIDEEIFSFYKILIKMRK 486
                        ||:|  : : |   ::  :| ::| || :||
Db    430 ----------------KQQDQSKIDWTLLGNDLNRSLFDYHKGLIGLRK 462


Search completed: July 15, 2002, 12:18:43
Job time: 135 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: July 15, 2002, 12:18:08 ; Search time 19.24 Seconds
(without alignments)
1124.960 Million cell updates/sec

Title: US-09-298-924-6
Perfect score: 2978
Sequence: 1 MTFAYKIDGNEVIFTLWAPY.....PQHIEEGKYEFDKGFALYKL 559

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1060.5	35.6	598	1	TREZ_ARTSP	Q44316 arthrobacte
2	1025	34.4	596	1	TREZ_RHISP	Q53238 rhizobium s
3	1024.5	34.4	575	1	TREZ_ARTRM	Q9ajn6 arthrobacte
4	997.5	33.5	580	1	TREZ_MYCTU	Q10769 mycobacteri
5	965	32.4	589	1	TREZ_BREHE	O52520 brevibacter
6	391	13.1	666	1	GLGB_BACCL	P30537 bacillus ca
7	374.5	12.6	627	1	GLGB_BACSU	P39118 bacillus su
8	374.5	12.6	639	1	GLGB_BACST	P30538 bacillus st
9	340	11.4	730	1	GLGB_HAEIN	P45177 haemophilus
10	335	11.2	764	1	GLGB_STRAU	P52980 streptomyce
11	333.5	11.2	770	1	GLGB_SYNY3	P52981 synechocyst
12	326.5	11.0	773	1	GLGB_SYNP7	P16954 synechococc
13	326	10.9	731	1	GLGB_MYCTU	Q10625 mycobacteri
14	321	10.8	639	1	GLGB_BUTFI	P30539 butyrivibri
15	318	10.7	659	1	GLGX_HAEIN	P45178 haemophilus
16	318	10.7	734	1	GLGB_AGRTU	P52979 agrobacteri
17	314	10.5	843	1	PULA_THEMA	O33840 thermotoga
18	303	10.2	561	1	TREC_BACSU	P39795 bacillus su
19	300	10.1	498	1	AMY3_DICTH	P14899 dictyoglomu
20	300	10.1	581	1	MAXS_YEAST	P40884 saccharomyc
21	299	10.0	520	1	AMY_BACME	P20845 bacillus me
22	298.5	10.0	562	1	O16G_BACTR	P29094 bacillus th
23	298	10.0	558	1	O16G_BACCE	P21332 bacillus ce
24	298	10.0	562	1	AMY2_DICTH	P14898 dictyoglomu
25	294.5	9.9	574	1	CDAS_THEET	P29964 thermoanaer
26	294.5	9.9	728	1	GLGB_ECOLI	P07762 escherichia
27	288.5	9.7	861	1	GLGB_SOLTU	P30924 solanum tub
28	287	9.6	657	1	GLGX_ECOLI	P15067 escherichia
29	285.5	9.6	702	1	GLGB_HUMAN	Q04446 homo sapien
30	283	9.5	963	1	TRES_THEAQ	O06458 thermus aqu
31	282.5	9.5	555	1	O16G_BACCO	Q45101 bacillus co
32	282.5	9.5	799	1	GLGB_MAIZE	Q08047 zea mays (m
33	278	9.3	820	1	GLGB_ORYSA	Q01401 oryza sativ
34	276	9.3	478	1	YDD1_SCHPO	Q10427 schizosacch
35	271.5	9.1	721	1	GLGX_MYCTU	Q10767 mycobacteri
36	266.5	8.9	588	1	NEPU_BACST	P38940 bacillus st
37	266	8.9	704	1	GLGB_YEAST	P32775 saccharomyc
38	265.5	8.9	586	1	AMYM_BACAD	P32818 bacillus ac
39	265	8.9	585	1	NEPU_THEVU	Q08751 thermoactin
40	258	8.7	1196	1	AMYB_PAEPO	P21543 paenibacill
41	257.5	8.6	508	1	O16G_BACSP	P29093 bacillus sp
42	257	8.6	494	1	AMYI_SACFI	P21567 saccharomyc
43	256.5	8.6	777	1	ISOA_FLASP	O32611 flavobacter
44	256	8.6	535	1	DEXB_STRPN	Q54796 streptococc
45	254	8.5	561	1	O16G_BACSU	O06994 bacillus su

ALIGNMENTS

RESULT 1
TREZ_ARTSP
ID TREZ_ARTSP STANDARD; PRT; 598 AA.
AC Q44316;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Malto-oligosyltrehalose trehalohydrolase (EC 3.2.1.141) (MTHase) (4-
DE alpha-D-{(1->4)-alpha-D-glucano}trehalose trehalohydrolase)
DE (Maltooligosyl trehalose trehalohydrolase).
GN TREZ.
OS Arthrobacter sp. (strain Q36).
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Micrococcineae; Micrococcaceae; Arthrobacter.
OX NCBI_TaxID=1667;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96195835; PubMed=8605217;
RA Maruta K., Hattori K., Nakada T., Kubota M., Sugimoto T., Kurimoto M.;
RT "Cloning and sequencing of trehalose biosynthesis genes from
RT Arthrobacter sp. Q36.";
RL Biochim. Biophys. Acta 1289:10-13(1996).
CC -!- CATALYTIC ACTIVITY: Hydrolysis of alpha-(1->4)-D-glucosidic
CC linkage in 4-alpha-D-{(1->4)-alpha-D-glucanosyl}n trehalose to
CC yield trehalose and alpha-(1->4)-D-glucan.
CC -!- PATHWAY: Trehalose biosynthesis.
CC -!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC ----------

CC ----------
DR EMBL; D63343; BAA09668.1; -.
DR InterPro; IPR000461; Alpha_amylase.
DR InterPro; IPR004193; isoamylase_N.
DR Pfam; PF00128; alpha-amylase; 1.
DR Pfam; PF02922; isoamylase_N; 1.
KW Hydrolase; Glycosidase.
FT ACT_SITE 267 267 BY SIMILARITY.
SQ SEQUENCE 598 AA; 65831 MW; 8B5C610AD3766947 CRC64;

Query Match 35.6%; Score 1060.5; DB 1; Length 598;
Best Local Similarity 39.7%; Pred. No. 4e-62;
Matches 232; Conservative 92; Mismatches 197; Indels 63; Gaps 16;

```
Qy       1 MTFAYKIDGNEVI-----FTLWAPYQKSVKLKVLEKG-LYEMER------DEKGYFTI-- 46
           ||  |  :  : :      : :|||  :|| |  |  |  | |:|      :: |::|
Db       1 MTHTYPREAAKPVLGPARYDVWAPNAESVTL--LAGGERYAMQRRAETGPEDAGWWTAAG 58

Qy      47 --TLNNVKVRDRYKYVLD-DASEIPDPASRYQPEGVHGPSQIIQESK-EFNNETFLKKE- 101
```

```
            |  ||     | |:|| | : :||| :| ||:|||  |:    |   : :: :  :|
Db     59 APTDGNVD----YGYLLDGDETPLPDPRTRRQPDGVHALSRTFDPSAYSWQDDAWQGREL 114

Qy    102 -DLIIYEIHVGTFTPEGTFEGVIRKLDYLKDLGITAIEIMPIAQFPGKRDWGYDGVYLYA 160
          :|||:|:|||||||| |    ||||| ||:  ||::|:  | |  :||||||  :|
Db    115 QGAVIYELHLGTFTPEGTLEAAAGKLDYLAGLGVDFIELLPVNAFNGTHNWGYDGVQWFA 174

Qy    161 VQNSYGGPEGFRKLVDEAHKKGLGVILDVVYNHVGPEGNYMVKLGPYFSQKYKTPWGLTF 220
          |  :||||| ::: || ||  ||||| ||||||:|| |||: : |||  |     || :
Db    175 VHEAYGGPEAYQRFVDAAHAAGLGVIQDVVYNHLGPSGNYLPRFGPYLKQGEGNTWGDSV 234

Qy    221 NFDDAESDEVRKFILENVEYWIKEYNVDGFRLDAVHAIIDTSPKHILEEIADVVHKYNRI 280
          | |   || ||::||:|:  |:::| ||| |||||||: |    ||||:   :  : :
Db    235 NLDGPGSDHVRRYILDNLAMWLRDYRVDGLRLDAVHALKDERAVHILEDFGALADQISAE 294

Qy    281 V------IAESDLNDPRVVNPKEKCGYNIDAQWVDDFHHSIHAYLTGERQGYYTDFGNLD 334
          |      |||||||:||:: |::  || :: || |||||::|  :|||  |||:|| :|
Db    295 VGRPLTLIAESDLNNPRLLYPRDVNGYGLEGQWSDDFHHAVHVNVTGETTGYYSDFDSLA 354

Qy    335 DIVKSYKDVFVYDGKYSNFRRKTHGEPVGELDGCNF--------VVYIQNHDQVGNRGKG 386
          : |  :| | :|| ||:|| : || |:      ||        ||  |||||:|||  |
Db    355 ALAKVLRDGFFHDGSYSSFRERHHGRPI------NFSAVHPAALVVCSQNHDQIGNRATG 408

Qy    387 ERIIKLVDRESYKIAAALYLLSPYIPMIFMGEEYGEENPFYFFSDFSDSKLIQGVREGRK 446
          :|: : :   | :|| | |  |: ||: |||||   |: ||:   : :| :   |||
Db    409 DRLSQTLPYGSLALAAVLTLTGPFTPMLLMGEEYGASTPWQFFTSHPEPELGKATAEGRI 468

Qy    447 KE---NGQD----TDPQDESTFNASKLSWKIDEE-----IFSFYKILIKMRKELSIAC-- 492
          ||    | |     |||| ||  ||| |    |     :   |: |  :|:
Db    469 KEFERMGWDPAVVPDPQDPETFRRSKLDWAEAAEGDHARLLELYRSLTALRRSTPDLTKL 528

Qy    493 ---DRRVNVVNGENWLIIKGREYFSLYVFSKSSIEVKYSGTLLL 533
            | :|       ||  :      |  ||:  : :  :|| ||
Db    529 GFEDTQVAFDEDARWLRFRRGGVQVLLNFSEQPVSLDGAGTALL 572


RESULT   2
TREZ_RHISP
ID   TREZ_RHISP     STANDARD;      PRT;   596 AA.
AC   Q53238;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Malto-oligosyltrehalose trehalohydrolase (EC 3.2.1.141) (MTHase) (4-
DE   alpha-D-{(1->4)-alpha-D-glucano}trehalose trehalohydrolase)
DE   (Maltooligosyl trehalose trehalohydrolase).
GN   TREZ.
OS   Rhizobium sp. (strain M-11).
OC   Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC   Rhizobiaceae; Rhizobium.
OX   NCBI_TaxID=391;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96219094; PubMed=8829547;
RA   Maruta K., Hattori K., Nakada T., Kubota M., Sugimoto T., Kurimoto M.;
RT   "Cloning and sequencing of trehalose biosynthesis genes from Rhizobium
RT   sp. M-11.";
RL   Biosci. Biotechnol. Biochem. 60:717-720(1996).
CC   -!- CATALYTIC ACTIVITY: Hydrolysis of alpha-(1->4)-D-glucosidic
CC       linkage in 4-alpha-D-{(1->4)-alpha-D-glucanosyl}(n) trehalose to
CC       yield trehalose and alpha-(1->4)-D-glucan.
CC   -!- PATHWAY: Trehalose biosynthesis.
CC   -!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC       KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D78001; BAA11187.1; -.
DR   InterPro; IPR000461; Alpha_amylase.
DR   Pfam; PF00128; alpha-amylase; 1.
KW   Hydrolase; Glycosidase.
FT   ACT_SITE    265    265       BY SIMILARITY.
SQ   SEQUENCE   596 AA;  65262 MW;  73EA80AE0534DDCD CRC64;


  Query Match             34.4%;  Score 1025;  DB 1;  Length 596;
  Best Local Similarity   40.3%;  Pred. No. 8.4e-60;
  Matches  229;  Conservative   80;  Mismatches  197;  Indels   62;  Gaps   16;

Qy     14 FTLWAPYQKSVKLKVLEKG-LYEMER------DEKGYFTITLNNVKVRDRYKYVLD-DAS 65
          | :|||   :| |  |  |  ||| |       ::|::|           | |:|| |
Db     17 FDIWAPEAGTVTL--LAGGERYEMGRRPGNGPADEGWWTAADAPTGADVDYGYLLDGDEI 74

Qy     66 EIPDPASRYQPEGVHGPSQIIQESKEFNNETFLKKEDL---IIYEIHVGTFTPEGTFEGV 122
          :||| :| ||||||  |:          :   :  :|   :|||:|:|||||||| :
Db     75 PLPDPRTRRQPEGVHALSRTFDPGAHRWQDAGWQGRELQGSVIYELHIGTFTPEGTLDAA 134

Qy    123 IRKLDYLKDLGITAIEIMPIAQFPGKRDWGYDGVYLYAVQNSYGGPEGFRKLVDEAHKKG 182
            |||||  |||  ||::|:  | |  :||||||  :||   ||||  ::: || ||  |
Db    135 AGKLDYLAGLGIDFIELLPVNAFNGTHNWGYDGVQWFAVHEGYGGPAAYQRFVDAAHAAG 194

Qy    183 LGVILDVVYNHVGPEGNYMVKLGPYFSQKYKTPWGLTFNFDDAESDEVRKFILENVEYWI 242
          |||| ||||||:|| |||: : |||         || : | |   || ||::||:||  |:
Db    195 LGVIQDVVYNHLGPSGNYLPRYGPYLKHGEGNTWGDSVNLDGPGSDHVRQYILDNVAMWL 254

Qy    243 KEYNVDGFRLDAVHAIIDTSPKHILEE---IADVVHKYN---RIVIAESDLNDPRVVNPK 296
          ::| ||| ||||||: |    |||||   :|| :          :|||||||:||:: |:
Db    255 RDYRVDGLRLDAVHALKDERAVHILEEFGALADALSSEGGRPLTLIAESDLNNPRLLYPR 314

Qy    297 EKCGYNIDAQWVDDFHHSIHAYLTGERQGYYTDFGNLDDIVKSYKDVFVYDGKYSNFRRK 356
          :  || :  || |||||::|  ::||  |||:|| :|  : |  :| | :|| ||:|| :
Db    315 DVNGYGLAGQWSDDFHHAVHVNVSGETTGYYSDFDSLGALAKVLRDGFFHDGSYSSFRGR 374

Qy    357 THGEPVGELDGCNF--------VVYIQNHDQVGNRGKGERIIKLVDRESYKIAAALYLLS 408
            || |:      ||         ||  |||||:|||  |:|: : :   |  :|| | |
Db    375 CHGRPI------NFSAVHPAALVVCSQNHDQIGNRATGDRLSQSLPYGSLALAAVLTLTG 428

Qy    409 PYIPMIFMGEEYGEENPFYFFSDFSDSKLIQGVREGRKKE---NGQD----TDPQDESTF 461
          |: ||:|||||||   |: ||:   : :| :   ||| :|    | |     ||||  ||
Db    429 PFTPMLFMGEEYGATTPWQFFTSHPEPELGKATAEGRIREFERMGWDPAVVPDPQDPETF 488

Qy    462 NASKLSWK-----IDEEIFSFYKILIKMRK------ELSIACDRRVNVVNGENWLIIKGR 510
            ||| |          :   |: || :|:       |  | | |   :   ||
Db    489 TRSKLDWAEASAGDHARLLELYRSLITLRRSTPELARLGFA-DTAVEFDDDARWL----- 542

Qy    511 EYFS-----LYVFSKSSIEVKYSGTLLL 533
            |:      :  |:   | :   || ||
Db    543 RYWRGGVQVVLNFADRPISLDRPGTALL 570


RESULT   3
TREZ_ARTRM
ID   TREZ_ARTRM     STANDARD;      PRT;   575 AA.
AC   Q9AJN6;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Malto-oligosyltrehalose trehalohydrolase (EC 3.2.1.141) (MTHase) (4-
DE   alpha-D-{(1->4)-alpha-D-glucano}trehalose trehalohydrolase)
DE   (Maltooligosyl trehalose trehalohydrolase).
GN   TREZ.
OS   Arthrobacter ramosus.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Micrococcineae; Micrococcaceae; Arthrobacter.
OX   NCBI_TaxID=1672;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S34;
RA   Yamamoto T., Maruta K., Watanabe H., Yamashita H., Kubota M.,
RA   Fukuda S., Kurimoto M.;
```

```
RT   "Trehalose producing operon treYZ from Arthrobacter ramosus S34.";
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- CATALYTIC ACTIVITY: Hydrolysis of alpha-(1->4)-D-glucosidic
CC       linkage in 4-alpha-D-[(1->4)-alpha-D-glucanosyl](n) trehalose to
CC       yield trehalose and alpha-(1->4)-D-glucan.
CC   -!- PATHWAY: Trehalose biosynthesis.
CC   -!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC       KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AB045141; BAB40766.1; -.
DR   InterPro; IPR000461; Alpha_amylase.
DR   Pfam; PF00128; alpha-amylase; 1.
KW   Hydrolase; Glycosidase.
FT   ACT_SITE    250    250       BY SIMILARITY.
SQ   SEQUENCE   575 AA;  63079 MW;  284A3F20207E228B CRC64;


  Query Match             34.4%;  Score 1024.5;  DB 1;  Length 575;
  Best Local Similarity   39.4%;  Pred. No. 8.6e-60;
  Matches  215;  Conservative   90;  Mismatches  189;  Indels   51;  Gaps   13;

Qy      14 FTLWAPYQKSVKLKVLEKGLYE--MERDEKGYFTITL---NNVKVRDRYKYVLDDASEIP 68
           | :|||    | | |: :|  |  : ||| |:: :         : | | |::|
Db       5 FPVWAPQAAQVTL-VVGQGRAELPLTRDENGWWALQQPWDGGPDLVD-YGYLVDGKGPFA 62

Qy      69 DPASRYQPEGVHGPSQIIQESKEFNNETFLKKED---------LIIYEIHVGTFTPEGTF 119
           || |  || |||       :  :||:   :   :|          :|||:||||||||||
Db      63 DPRSLRQPRGVH------ELGREFDPARYAWGDDGWRGRDLTGAVIYELHVGTFTPEGTL 116

Qy     120 EGVIRKLDYLKDLGITAIEIMPIAQFPGKRDWGYDGVYLYAVQNSYGGPEGFRKLVDEAH 179
           :  ||:||:|  ||: |:|::|:  | |   ||||||  |||   ||||| ::: ||  |
Db     117 DSAIRRLDHLVRLGVDAVELLPVNAFNGTHGWGYDGVLWYAVHEPYGGPEAYQRFVDACH 176

Qy     180 KKGLGVILDVVYNHVGPEGNYMVKLGPYFSQKYKTPWGLTFNFDDAESDEVRKFILENVE 239
            :|| |: ||||||:|| ||::   |||          ||   | |   |||||::|::|
Db     177 ARGLAVVQDVVYNHLGPSGNHLPDFGPYLGSGAANTWGDALNLDGPLSDEVRRYIIDNAV 236

Qy     240 YWIKEYNVDGFRLDAVHAIIDTSPKHILEEIA----DVVHKYNR--IVIAESDLNDPRVV 293
           ||::: : || |||||||: |    |:|||:|    ::  :  |   :|||||||||:::
Db     237 YWLRDMHADGLRLDAVHALRDARALHLLEELAARVDELAGELGRPLTLIAESDLNDPKLI 296

Qy     294 NPKEKCGYNIDAQWVDDFHHSIHAYLTGERQGYYTDFGNLDDIVKSYKDVFVYDGKYSNF 353
             :   || :|||| || ||:|| :|||  ||| |||  |  :|| ::  : :|| :|:|
Db     297 RSRAAHGYGLDAQWDDDVHHAVHANVTGETVGYYADFGGLGALVKVFQRGWFHDGTWSSF 356

Qy     354 RRKTHGEPVG-ELDGCNFVVYIQNHDQVGNRGKGERIIKLVDRESYKIAAALYLLSPYIP 412
           | : || |:  ::     | : |:|||||||  |:|:   |   |   |||| || |: |
Db     357 RERHHGRPLDPDIPFRRLVAFAQDHDQVGNRAVGDRMSAQVGEGSLAAAAALVLLGPFTP 416

Qy     413 MIFMGEEYGEENPFYFFSDFSDSKLIQGVREGRKKE---NGQD----TDPQDESTFNASK 465
           |:|||||:|   |: ||:   : :| :    ||  |    | |     |||| :||  |
Db     417 MLFMGEEWGARTPWQFFTSHPEPELGEATARGRIAEFARMGWDPAVVPDPQDPATFARSH 476

Qy     466 LSWKIDEE-----IFSFYKILIKMRKELSIACDRRVNVVNGENWLIIKGREYFSLYVFSK 520
           | |   |      : :||  || :|:|| :    |           :   |   :: ||:
Db     477 LDWSEPEREPHAGLLAFYTDLIALRRELPVDAPARE----------VDADEARGVFAFSR 526

Qy     521 SSIEV 525
             : |
Db     527 GPLRV 531


RESULT   4
TREZ_MYCTU
ID   TREZ_MYCTU     STANDARD;      PRT;   580 AA.
AC   Q10769;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Malto-oligosyltrehalose trehalohydrolase (EC 3.2.1.141) (MTHase) (4-
DE   alpha-D-[(1->4)-alpha-D-glucano]trehalose trehalohydrolase)
DE   (Maltooligosyl trehalose trehalohydrolase).
GN   TREZ OR RV1562C OR MT1613 OR MTCY48.03.
OS   Mycobacterium tuberculosis.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX   NCBI_TaxID=1773;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=H37RV;
RX   MEDLINE=98295987; PubMed=9634230;
RA   Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA   Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekaia F.,
RA   Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA   Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA   Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA   Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA   Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA   Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT   "Deciphering the biology of Mycobacterium tuberculosis from the
RT   complete genome sequence.";
RL   Nature 393:537-544(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CDC 1551 / Oshkosh;
RA   Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA   Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA   Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA   Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA   Bishai W.;
RT   "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT   laboratory strains.";
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- CATALYTIC ACTIVITY: Hydrolysis of alpha-(1->4)-D-glucosidic
CC       linkage in 4-alpha-D-[(1->4)-alpha-D-glucanosyl](n) trehalose to
CC       yield trehalose and alpha-(1->4)-D-glucan.
CC   -!- PATHWAY: Trehalose biosynthesis.
CC   -!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC       KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z74020; CAA98329.1; -.
DR   EMBL; AE007027; AAK45880.1; ALT_INIT.
DR   TIGR; MT1613; -.
DR   TubercuList; Rv1562c; -.
DR   InterPro; IPR000461; Alpha_amylase.
DR   InterPro; IPR004193; isoamylase_N.
DR   Pfam; PF00128; alpha-amylase; 1.
DR   Pfam; PF02922; isoamylase_N; 1.
KW   Hydrolase; Glycosidase; Complete proteome.
FT   ACT_SITE    247    247       BY SIMILARITY.
SQ   SEQUENCE   580 AA;  64076 MW;  506190468F44B862 CRC64;


  Query Match             33.5%;  Score 997.5;  DB 1;  Length 580;
  Best Local Similarity   38.3%;  Pred. No. 5.1e-58;
  Matches  212;  Conservative   91;  Mismatches  215;  Indels   35;  Gaps   11;

Qy      14 FTLWAPYQKSVKLKVLEKGLYEMERDEKGYFTITLNNVKVRDRYKYVL-DDASEIPDPAS 72
           | :|||    |:| |    :: | |   |::  |:       || |:| || : :||| |
Db       4 FRVWAPKPALVRLDV-NGAVHAMTRSADGWWHTTV-AAPADARYGYLLDDDPTVLPDPRS 61
```

```
Qy      73 RYQPEGVHGPSQIIQESKEF---NNET---FLKKEDLIIYEIHVGTFTPEGTFEGVIRKL 126
           ||:|||  ||  :   :|     :|       |  :|||:|:|||| |||:  | ||
Db      62 ARQPDGVHARSQRWEPPGQFGAARTDTGWPGRSVEGAVIYELHIGTFTTAGTFDAAIEKL 121

Qy     127 DYLKDLGITAIEIMPIAQFPGKRDWGYDGVYLYAVQNSYGGPEGFRKLVDEAHKKGLGVI 186
           ||| |||| :|:||:  | | | |||||  |:|  ||||:|  : :|  | : |||:
Db     122 DYLVDLGIDFVELMPVNSFAGTRGWGYDGVLWYSVHEPYGGPDGLVRFIDACHARRLGVL 181

Qy     187 LDVVYNHVGPEGNYMVKLGPYFSQKYKTPWGLTFNFDDAESDEVRKFILENVEYWIKEYN 246
           :| |:||:|| |||: : ||| |     |||   |   |:|||| :|::    |:::::
Db     182 IDAVFNHLGPSGNYLPRFGPYLSSA-SNPWGDGINIAGADSDEVRHYIIDCALRWMRDFH 240

Qy     247 VDGFRLDAVHAIIDTSPKHILEEIADVVHKYN------RIVIAESDLNDPRVVNPKEKCG 300
           || |||||||::||:  |:|||:|:     :        :|||:| ||||::       |
Db     241 ADGLRLDAVHALVDTTAVHVLEELANATRWLSGQLGRPLSLIAETDRNDPRLITRPSHGG 300

Qy     301 YNIDAQWVDDFHHSIHAYLTGERQGYYTDFGNLDDIVKSYKDVFVYDGKYSNFRRKTHGE 360
           | | ||| || ||:||  ::|||||| |||:|  :  : :: : :  | ||:|||: ||
Db     301 YGITAQWNDDIHHAIHTAVSGERQGYYADFGSLATLAYTLRNGYFHAGTYSSFRRRRHGR 360

Qy     361 PV--GELDGCNFVVYIQNHDQVGNRGKGERIIKLVDRESYKIAAALYLLSPYIPMIFMGE 418
           :   :     : |   |||||||  |:|  : :       | ||| | |||  |:||||
Db     361 ALDTSAIPATRLLAYTCTHDQVGNRALGDRPSQYLTGGQLAIKAALTLGSPYTAMLFMGE 420

Qy     419 EYGEENPFYFFSDFSDSKLIQGVREGRKKENGQD-------TDPQDESTFNASKLSWK-- 469
           |:|  :|| ||    : :|      |||:|  :         ||||  ||   ||:|
Db     421 EWGASSPFQFFCSHPEPELAHSTVAGRKEEFAEHGWAADDIPDPQDPQTFQRCKLNWAEA 480

Qy     470 ---IDEEIFSFYKILIKMRKELSIACD-----RRVNVVNGENWLIIKGREYFSLYVFSKS 521
              :  ||: || :|   :   |       |:    : |:::: :
Db     481 GSGEHARLHRFYRDLIALRHNEADLADPWLDHLMVDYDEQQRWVVMRRGQLMIACNLGAE 540

Qy     522 SIEVKYSGTLLLS 534
             |  || |:|:
Db     541 PTCVPVSGELVLA 553


RESULT   5
TREZ_BREHE
ID   TREZ_BREHE     STANDARD;      PRT;   589 AA.
AC   O52520;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Malto-oligosyltrehalose trehalohydrolase (EC 3.2.1.141) (MTHase) (4-
DE   alpha-D-[(1->4)-alpha-D-glucano]trehalose trehalohydrolase)
DE   (Maltooligosyl trehalose trehalohydrolase).
GN   TREZ.
OS   Brevibacterium helvolum.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Micrococcineae; Brevibacteriaceae; Brevibacterium.
OX   NCBI_TaxID=1704;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 11822;
RA   Kwon T., Kim C.H., Choi Y.D.;
RT   "Maltooligosyl trehalose synthase/trehalohydrolase encode proteins for
RT   trehalose production in Brevibacterium helvolum.";
RL   Submitted (DEC-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- CATALYTIC ACTIVITY: Hydrolysis of alpha-(1->4)-D-glucosidic
CC       linkage in 4-alpha-D-[(1->4)-alpha-D-glucanosyl](n) trehalose to
CC       yield trehalose and alpha-(1->4)-D-glucan.
CC   -!- PATHWAY: Trehalose biosynthesis.
CC   -!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC       KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF039919; AAB95369.1; -.
DR   InterPro; IPR000461; Alpha_amylase.
DR   Pfam; PF00128; alpha-amylase; 1.
KW   Hydrolase; Glycosidase.
FT   ACT_SITE    258    258       BY SIMILARITY.
SQ   SEQUENCE   589 AA;  64217 MW;  C282314B6E9BA029 CRC64;


  Query Match             32.4%;  Score 965;  DB 1;  Length 589;
  Best Local Similarity   38.5%;  Pred. No. 6.9e-56;
  Matches  218;  Conservative   88;  Mismatches  216;  Indels   44;  Gaps   13;

Qy      14 FTLWAPYQKSVKLKVLEKGLYEMERDE-----KGYFTITLNNVKVRDRYKYVLD-DASEI 67
           | :|||   || | | :   | |:: |     :|::| :          | |:|| : : :
Db      11 FDVWAPDVSSVVL-VADGRQYPMQKKETAPGSEGWWTASDAPPNGDVDYGYLLDGNTTPV 69

Qy      68 PDPASRYQPEGVHGPSQII-------QESKEFNNETFLKKEDLIIYEIHVGTFTPEGTFE 120
           |:| ||  | |||  |:         |:|:    |     :  :||::|||| ||:|| :
Db      70 PEPRSRRLPAGVHNHSRTYNPPPYRWQDSRWRGKEL----QGTLIYQLHVGTSTPDGTLD 125

Qy     121 GVIRKLDYLKDLGITAIEIMPIAQFPGKRDWGYDGVYLYAVQNSYGGPEGFRKLVDEAHK 180
              || || |||  ||::|:  | |  :||||||  | |   ||||  ::: || ||
Db     126 AAGEKLSYLVDLGIDFIELLPVNGFNGTHNWGYDGVQWYTVHEGYGGPAAYQRFVDAAHA 185

Qy     181 KGLGVILDVVYNHVGPEGNYMVKLGPYFSQKYKTPWGLTFNFDDAESDEVRKFILENVEY 240
           |||||  ||||||:|  |||  ||||   |      | : | | | ||  |::||:|
Db     186 AGLGVIQDVVYNHLGLRGNYFPKLGPNLKQGDANTLGDSVNLDGAGSDVFREYILDNAAL 245

Qy     241 WIKEYNVDGFRLDAVHAIIDTSPKHILEE---IADVVHKYN---RIVIAESDLNDPRVVN 294
           |: :|:|||   |||||: |    ||||:   : | :       : :|||||  |:||::
Db     246 WVGDYHVDGVGFDAVHAVRDERAVHILEDLGALGDAISGETGLPKTLIAESDFNNPRLIY 305

Qy     295 PKEKCGYNIDAQWVDDFHHSIHAYLTGERQGYYTDFGNLDDIVKSYKDVFVYDGKYSNFR 354
           |::  || :  || |||| ::|  ::||  |||:|| :|  : |  || |::|| ||:||
Db     306 PRDVNGYGLAGQWSDDFHTAVHVSVSGETTGYYSDFESLAVLAKVLKDGFLHDGSYSSFR 365

Qy     355 RKTHGEPVGE--LDGCNFVVYIQNHDQVGNRGKGERIIKLVDRESYKIAAALYLLSPYIP 412
           : || |:     :    ||  |||||:|||  |:|: : :       :|| | | ||: |
Db     366 GRHHGRPINPSLANPAALVVCNQNHDQIGNRATGDRLSQSLSYGQLAVAAVLTLTSPFTP 425

Qy     413 MIFMGEEYGEENPFYFFSDFSDSKLIQGVREGRKKE---NGQD----TDPQDESTFNASK 465
           |:|||||   |: ||:   : :| :   ||| ||    | |     ||||  ||| ||
Db     426 MLFMGEEYGASTPWQFFTSHPEPELGKATAEGRIKEFERMGWDPAVVPDPQDPETFNRSK 485

Qy     466 LSWK-----IDEEIFSFYKILIKMRKELSIACD-----RRVNVVNGENWLIIKGREYFSL 515
           | |          :   || |  :|:|     |       |:  :   ||  :      |
Db     486 LDWSEASTGDHARLLELYKSLTALRREHPDLADLGFGQTEVSFDDDAGWLRFRPVSVEVL 545

Qy     516 YVFSKSSIEV-KYSGTLLLSSNNSFP 540
              | : :   :| |||:::   |
Db     546 VNLSDAKVRLDDAAGDLLLATDEGNP 571


RESULT   6
GLGB_BACCL
ID   GLGB_BACCL     STANDARD;      PRT;   666 AA.
AC   P30537;
DT   01-APR-1993 (Rel. 25, Created)
DT   01-APR-1993 (Rel. 25, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   1,4-alpha-glucan branching enzyme (EC 2.4.1.18) (Glycogen branching
DE   enzyme).
GN   GLGB.
OS   Bacillus caldolyticus.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Bacillus/Staphylococcus group; Bacillus.
OX   NCBI_TaxID=1394;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93208370; PubMed=1296817;
RA   Kiel J.A.K.W., Boels J.M., Beldman G., Venema G.;
```

```
RT   "The glgB gene from the thermophile Bacillus caldolyticus encodes a
RT   thermolabile branching enzyme.";
RL   DNA Seq. 3:221-232(1992).
CC   -!- CATALYTIC ACTIVITY: Formation of 1,6-glucosidic linkages of
CC       glycogen.
CC   -!- PATHWAY: THIRD STEP IN GLYCOGEN BIOSYNTHESIS.
CC   -!- MISCELLANEOUS: OPTIMAL ACTIVITY AT APPROXIMATELY 39 DEGREES
CC       CELSIUS.
CC   -!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC       KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z14057; CAA78440.1; -.
DR   InterPro; IPR000461; Alpha_amylase.
DR   InterPro; IPR004193; isoamylase_N.
DR   Pfam; PF00128; alpha-amylase; 1.
DR   Pfam; PF02922; isoamylase_N; 1.
KW   Glycogen biosynthesis; Transferase; Glycosyltransferase.
FT   ACT_SITE    309    309       BY SIMILARITY.
FT   ACT_SITE    352    352       BY SIMILARITY.
FT   ACT_SITE    420    420       BY SIMILARITY.
SQ   SEQUENCE   666 AA;  78096 MW;  1F64C31495BB1B6E CRC64;


  Query Match             13.1%;  Score 391;  DB 1;  Length 666;
  Best Local Similarity   24.8%;  Pred. No. 2.7e-18;
  Matches  155;  Conservative   99;  Mismatches  202;  Indels  170;  Gaps   31;

Qy       4 AYKIDGNEVI---------FTLWAPYQKSVKL--------------KVLEKGLYEMERD 39
           :|:: |  ||         | :|||: : |:|               ||       :
Db      22 SYELFGAHVIRGGGAVGTRFCVWAPHAREVRLVGSFNDWNGTNSPLTKV---------N 71

Qy      40 EKGYFTITLNNVKVRDRYKY--VLDDASEI--PDPASRYQPEGVHGPSQIIQESK--EFN 93
           ::| :|| :         |||  :  |   :   || : |      |  | |: : |  |:|
Db      72 DEGVWTIVVPENLEGHLYKYEIITPDGRVLLKADPYAFYSELRPHTAS-IVYDLKGYEWN 130

Qy      94 NETFLKK--------EDLIIYEIHVGTF--TPEGTF----EGVIRKLDYLKDLGITAIEI 139
           :  : :|         : ::|||:| |::   |:| |    |     : |: : | | ||:
Db     131 DSPWQRKKRRKRIYDQPMVIYELHFGSWKKKPDGRFYTYREMADELIPYVLERGFTHIEL 190

Qy     140 MPIAQFPGKRDWGYDGVYLYAVQNSYGGPEGFRKLVDEAHKKGLGVILDVVYNHV--GPE 197
           :|: : |  | |||  |   |:| : || |  |   ||  |: ||||:| |  |
Db     191 LPLVEHPLDRSWGYQGTGYYSVTSRYGTPHDFMYFVDRCHQAGLGVIIDWVPGHFCKDAH 250

Qy     198 GNYMVKLGP---YFSQKYKT--PWGLTFNFDDAESDEVRKFILENVEYWIKEYNVDGFRL 252
           | ||    |   | ::| :    || | || |    ||| |:: |  :|:: |:|||||:
Db     251 GLYMFDGAPTYEYANEKDRENYVWG-TANF-DLGKPEVRSFLISNALFWLEYYHVDGFRV 308

Qy     253 DAVHAII----------DTSPKHILEEIADVVHKY--NRIVIAESDLNDPRVVNPKEKCG 300
           |||  ::          :      | :: : |  |  |  :|||   : |||  |    |
Db     309 DAVANMLYWPNNDRLYENPYAVEFLRQLNEAVFAYDPNVWMIAEDSTDWPRVTAPTYDGG 368

Qy     301 YNIDAQWVDDFHHSIHAYLTGERQGYYTDFGNLDDIVKSYKDVFVYDGKYSNFRRKTHGE 360
           : :|                  : | ::|::| | :   :: ||::
Db     369 LGFNYKW--------------------NMGWMNDMLK-YMETPPHERKYAH-------- 398

Qy     361 PVGELDGCNFVVYIQN------HDQVGNRGKGERIIKLVDRESYKIAAALYLLSPYI--- 411
             ::       | :|      ||:| : ||   : |:      |  | | ||  |:
Db     399 --NQVSFSLLYAYSENFILPFSHDEVVH-GKKSLLNKMPGSYEEKF-AQLRLLYGYMMAH 454

Qy     412 ---PMIFMGEEYGEENPFYFFSDFSDSKLIQGVREGRKKENGQDTDPQDESTFNASKLSW 468
               ::||| |         |: | : |                          | :| |
Db     455 PGKKLLFMGSE---------FAQFDEWKF------------------------AEELDW 480

Qy     469 KIDEEIFSFYKILIKMRKELSIACDRRVNVV-------NGENWLIIKGREYFSLYVFSKS 521
             : :  |  :: : :  |:| ||| :|           |  |: :   |  |:: | :

Db     481 VLFD--FELHRKMDEYVKQL-IACYKRYKPFYELDHDPRGFEWIDVHNAEQ-SIFSFIRR 536

Qy     522 SIEVKYSGTLLLSSNNSFPQHIEEGK 547
               | | :|:   |   |  :: |
Db     537 G---KKEGDVLVIVCNFTNQAYDDYK 559


RESULT   7
GLGB_BACSU
ID   GLGB_BACSU     STANDARD;      PRT;   627 AA.
AC   P39118;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   1,4-alpha-glucan branching enzyme (EC 2.4.1.18) (Glycogen branching
DE   enzyme).
GN   GLGB.
OS   Bacillus subtilis.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Bacillus/Staphylococcus group; Bacillus.
OX   NCBI_TaxID=1423;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168;
RX   MEDLINE=94195107; PubMed=8145641;
RA   Kiel J.A.K.W., Boels J.M., Beldman G., Venema G.;
RT   "Glycogen in Bacillus subtilis: molecular characterization of an
RT   operon encoding enzymes involved in glycogen biosynthesis and
RT   degradation.";
RL   Mol. Microbiol. 11:203-218(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98048467; PubMed=9387221;
RA   Lapidus A., Galleron N., Sorokin A., Ehrlich S.D.;
RT   "Sequencing and functional annotation of the Bacillus subtilis genes
RT   in the 200 kb rrnB-dnaB region.";
RL   Microbiology 143:3431-3441(1997).
CC   -!- FUNCTION: CATALYSES THE FORMATION OF THE ALPHA-1,6-GLUCOSIDIC
CC       LINKAGES IN GLYCOGEN BY SCISSION OF A 1,4-ALPHA-LINKED
CC       OLIGOSACCHARIDE FROM GROWING ALPHA-1,4-GLUCAN CHAINS AND THE
CC       SUBSEQUENT TRANSFER OF THE OLIGOSACCHARIDE TO OTHER PARTS OF
CC       ALPHA-1,4-GLUCAN CHAINS.
CC   -!- CATALYTIC ACTIVITY: Formation of 1,6-glucosidic linkages of
CC       glycogen.
CC   -!- PATHWAY: THIRD STEP IN GLYCOGEN BIOSYNTHESIS.
CC   -!- INDUCTION: EXPRESSED EXCLUSIVELY ON MEDIA CONTAINING CARBON
CC       SOURCES THAT ALLOW EFFICIENT SPORULATION.
CC   -!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC       KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z25795; CAA81040.1; -.
DR   EMBL; AF008220; AAC00214.1; -.
DR   EMBL; Z99119; CAB15076.1; -.
DR   PIR; S36624; S36624.
DR   SubtiList; BG10907; glgB.
DR   InterPro; IPR000461; Alpha_amylase.
DR   InterPro; IPR004193; isoamylase_N.
DR   Pfam; PF00128; alpha-amylase; 1.
DR   Pfam; PF02922; isoamylase_N; 1.
KW   Glycogen biosynthesis; Transferase; Glycosyltransferase;
KW   Complete proteome.
FT   ACT_SITE    309    309       BY SIMILARITY.
FT   ACT_SITE    352    352       BY SIMILARITY.
FT   ACT_SITE    420    420       BY SIMILARITY.
SQ   SEQUENCE   627 AA;  73665 MW;  64B0A5553B6767BA CRC64;
```

```
  Query Match             12.6%;  Score 374.5;  DB 1;  Length 627;
  Best Local Similarity   23.1%;  Pred. No. 3e-17;
  Matches  139;  Conservative   94;  Mismatches  223;  Indels  147;  Gaps   24;

Qy      14 FTLWAPYQKSVKLKVLEKGLYE--------MER-DEKGYFTITLNNVKVRDRYKY-VLDD 63
           | :|||:   |::     | :          | | :: | :|: :  :  ::|||| :: :
Db      41 FCVWAPHASEVRV----AGDFNSWSGEEHVMHRVNDNGIWTLFIPGIGEKERYKYEIVTN 96

Qy      64 ASEI--------------PDPAS-RYQPEGVHGPSQIIQESKEFNNETFLKKEDLIIYEI 108
             ||              |: ||  |  ||     |  |: ::      | :: : |||:
Db      97 NGEIRLKADPYAIYSEVRPNTASLTYDLEGYSWQDQKWQKKQKAKT---LYEKPVFIYEL 153

Qy     109 HVGTFTPEGTFEGVIRK------LDYLKDLGITAIEIMPIAQFPGKRDWGYDGVYLYAVQ 162
           |:|::          |      : |:|  | | ||::|: : |  | ||| |    |:
Db     154 HLGSWKKHSDGRHYSYKELSQTLIPYIKKHGFTHIELLPVYEHPYDRSWGYQGTGYYSPT 213

Qy     163 NSYGGPEGFRKLVDEAHKKGLGVILDVVYNHV--GPEGNYMVKLGPYFSQKYKTP----- 215
           : :| |    | ||| |:: :|||| |  |      | ||    | :  | :
Db     214 SRFGPPHDLMKFVDECHQQNIGVILDWVPGHFCKDAHGLYMFDGEPLYEYKEERDRENWL 273

Qy     216 WGLTFNFDDAESDEVRKFILENVEYWIKEYNVDGFRLDAVHAII---DTSPKHILEEIAD 272
           || | || |    ||  |:: |  || : |::||||:||| |:   :   :|      |
Db     274 WG-TANF-DLGKPEVHSFLISNALYWAEFYHIDGFRVDAVANILYWPNQDERHTNPYAVD 331

Qy     273 VVHKYNR---------IVIAESDLNDPRVVNPKEKCGYNIDAQWVDDFHHSIHAYLTGER 323
           : | |:         ::|||     |:|    |: |      :|
Db     332 FLKKLNQTMREAYPHVMMIAEDSTEWPQVTGAVEEGGLGFHYKW---------------- 375

Qy     324 QGYYTDFGNLDDIVKSYKDVFVYDGKYSNFRRKTHGEPVGELDGCNFVVYIQ-------- 375
           : | ::|::| | :                  |  |    |: ::
Db     376 -----NMGWMNDVLK-YMET----------------PPEERRHCHQLISFSLLYAFSEH 412

Qy     376 -----NHDQVG-------NRGKGERIIKLVDRESYKIAAALYLLSPYIPMIFMGEEYGEE 423
                :||:|        |:  |:   |      |::      : |   :|||| |: :
Db     413 FVLPFSHDEVVYGKKSLLNKMPGDYWQKFA---QYRLLLGYMTVHPGKKLIFMGSEFAQ- 468

Qy     424 NPFYFFSDFSDSKLIQGVREGRKKENGQDTDPQD------------ESTFNASKLSW--- 468
             | :: |:: :    :              ||              |   |    |
Db     469 -----FDEWKDTEQLDWFLDSFPMHQKASVFTQDLLRFYQKSKILYEHDHRAQSFEWIDV 523

Qy     469 -KIDEEIFSFYKILIKMRKELSIACDRRVNVVNGENWLIIKGREYFSLYVFSKSSIEVKY 527
            :: |||| :   |  : | | |:    | :  :     |  :|: |:   :|    |
Db     524 HNDEQSIFSFIRYGQKHGEALVIICNFTPVVYHQYD----VGVPFFTQYIEVLNSDSETY 579

Qy     528 SGT 530
             |:
Db     580 GGS 582


RESULT   8
GLGB_BACST
ID   GLGB_BACST     STANDARD;      PRT;   639 AA.
AC   P30538;
DT   01-APR-1993 (Rel. 25, Created)
DT   01-APR-1993 (Rel. 25, Last sequence update)
DT   01-FEB-1995 (Rel. 31, Last annotation update)
DE   1,4-alpha-glucan branching enzyme (EC 2.4.1.18) (Glycogen branching
DE   enzyme).
GN   GLGB.
OS   Bacillus stearothermophilus.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Bacillus/Staphylococcus group; Geobacillus.
OX   NCBI_TaxID=1422;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=1503-4R;
RX   MEDLINE=92079888; PubMed=1745226;
RA   Kiel J.A.K.W., Boels J.M., Beldman G., Venema G.;
RT   "Molecular cloning and nucleotide sequence of the glycogen branching
RT   enzyme gene (glgB) from Bacillus stearothermophilus and expression in
RT   Escherichia coli and Bacillus subtilis.";
RL   Mol. Gen. Genet. 230:136-144(1991).
CC   -!- CATALYTIC ACTIVITY: Formation of 1,6-glucosidic linkages of
CC       glycogen.
CC   -!- PATHWAY: THIRD STEP IN GLYCOGEN BIOSYNTHESIS.
CC   -!- MISCELLANEOUS: THE TEMPERATURE FOR OPTIMAL ACTIVITY IS
CC       APPROXIMATELY 55 DEGREES CELSIUS.
CC   -!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC       KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M35089; AAA22482.1; -.
DR   PIR; S18599; S18599.
DR   InterPro; IPR000461; Alpha_amylase.
DR   InterPro; IPR004193; isoamylase_N.
DR   Pfam; PF00128; alpha-amylase; 1.
DR   Pfam; PF02922; isoamylase_N; 1.
KW   Glycogen biosynthesis; Transferase; Glycosyltransferase.
FT   ACT_SITE    309    309       BY SIMILARITY.
FT   ACT_SITE    352    352       BY SIMILARITY.
FT   ACT_SITE    420    420       BY SIMILARITY.
SQ   SEQUENCE   639 AA;  74795 MW;  F5BD4446B371E03A CRC64;


  Query Match             12.6%;  Score 374.5;  DB 1;  Length 639;
  Best Local Similarity   23.6%;  Pred. No. 3e-17;
  Matches  141;  Conservative   99;  Mismatches  190;  Indels  167;  Gaps   29;

Qy       4 AYKIDGNEVI---------FTLWAPYQKSVKL--------------KVLEKGLYEMERD 39
           :|:: |  ||         | :|||: : |:|               ||  :|:: :
Db      22 SYELFGAHVIKKNGMVGTRFCVWAPHAREVRLVGSFNEWNGTNFNLMKVSNQGVWMIFIP 81

Qy      40 E--KGYFTITLNNVKVRDRYKYVL--DDASEI--PDPASRYQPEGVHGPSQIIQ-ESKEF 92
           |  :|:           |||  :  :| : :   || : |     |  | :  :  ::
Db      82 ENLEGHL-----------YKYEITTNDGNVLLKSDPYAFYSELRPHTASIVYNIKGYQW 129

Qy      93 NNETFLKK--------EDLIIYEIHVGTF--TPEGTF----EGVIRKLDYLKDLGITAIE 138
           |::|: :|        : | |||:| |::    :|:|     |      : |: : | | ||
Db     130 NDQTWRRKKQRKRIYDQPLFIYELHFGSWKKKEDGSFYTYQEMAEELIPYVLEHGFTHIE 189

Qy     139 IMPIAQFPGKRDWGYDGVYLYAVQNSYGGPEGFRKLVDEAHKKGLGVILDVVYNHV--GP 196
           ::|: : |  | ||| |:  |:  : || |      :|  |: |:|||| |  |
Db     190 LLPLVEHPFDRSWGYQGIGYYSATSRYGTPHDLMYFIDRCHQAGIGVILDWVPGHFCKDS 249

Qy     197 EGNYMVKLGPYF-----SQKYKTPWGLTFNFDDAESDEVRKFILENVEYWIKEYNVDGFR 251
             | ||    | :      :    || | || |    ||| |:: |  :|:: ::|||||
Db     250 HGLYMFDGAPAYEYANMQDRENYVWG-TANF-DLGKPEVRSFLISNALFWMEYFHVDGFR 307

Qy     252 LDAVHAII----------DTSPKHILEEIADVVHKY--NRIVIAESDLNDPRVVNPKEKC 299
           :|||  ::           :|     |::: : |  |  | ::|||   : |||  |
Db     308 VDAVANMLYWPNSDVLYKNTYAVEFLQKLNETVFAYDPNILMIAEDSTDWPRVTAPTYDG 367

Qy     300 GYNIDAQWVDDFHHSIHAYLTGERQGYYTDFGNLDDIVKSYKDVFVYDGKYSNFRRKTHG 359
           |   : :|                      : | ::||:        |       |:  |
Db     368 GLGFNYKW--------------------NMGWMNDIL-------TYMETPPEHRKYVHN 399

Qy     360 EPVGELDGCNFVVYIQN------HDQVGNRGKGERIIKLVDRESYKIAAAL----YLLS- 408
           :   '|       | :|      ||:| : ||   : |:      | |       |||:
Db     400 KVTFSL----LYAYSENFILPFSHDEVVH-GKKSLLSKMPGTYEEKFAQLRLLYGYLLTH 454

Qy     409 PYIPMIFMGEEYGEENPFYFFSDFSDSKLIQGVREGRKKENGQDTDPQDESTFNASKLSW 468
           |   ::||| |:|:     | :: |                              :| |
Db     455 PGKKLLFMGGEFGQ------FDEWKD----------------------------LEQLDW 480

Qy     469 KIDEEIFSFYKILIKMRKELSIACDRRVNVV-------NGENWLIIKGREYFSLYVF 518
             : :  |  :: :    ||| : | :|   :       :|  |: :   |  |:: |
Db     481 MLFD--FDMHRNMNMYVKEL-LKCYKRYKPLYELDHSPDGFEWIDVHNAEQ-SIFSF 533
```

```
RESULT   9
GLGB_HAEIN
ID   GLGB_HAEIN     STANDARD;      PRT;   730 AA.
AC   P45177;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   1,4-alpha-glucan branching enzyme (EC 2.4.1.18) (Glycogen branching
DE   enzyme).
GN   GLGB OR HI1357.
OS   Haemophilus influenzae.
OC   Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC   Haemophilus.
OX   NCBI_TaxID=727;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=RD / KW20 / ATCC 51907;
RX   MEDLINE=95350630; PubMed=7542800;
RA   Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA   Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA   McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA   Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA   Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA   Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA   Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA   Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA   Venter J.C.;
RT   "Whole-genome random sequencing and assembly of Haemophilus
RT   influenzae Rd.";
RL   Science 269:496-512(1995).
CC   -!- CATALYTIC ACTIVITY: Formation of 1,6-glucosidic linkages of
CC       glycogen.
CC   -!- PATHWAY: THIRD STEP IN GLYCOGEN BIOSYNTHESIS.
CC   -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC       KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U32815; AAC23004.1; -.
DR   TIGR; HI1357; -.
DR   InterPro; IPR000461; Alpha_amylase.
DR   InterPro; IPR004193; isoamylase_N.
DR   Pfam; PF00128; alpha-amylase; 1.
DR   Pfam; PF02922; isoamylase_N; 1.
KW   Glycogen biosynthesis; Transferase; Glycosyltransferase;
KW   Complete proteome.
FT   ACT_SITE    405    405       BY SIMILARITY.
FT   ACT_SITE    458    458       BY SIMILARITY.
FT   ACT_SITE    526    526       BY SIMILARITY.
SQ   SEQUENCE   730 AA;  83820 MW;  5B9575317F53769A CRC64;


  Query Match             11.4%;  Score 340;  DB 1;  Length 730;
  Best Local Similarity   24.1%;  Pred. No. 6.5e-15;
  Matches  150;  Conservative  101;  Mismatches  214;  Indels  158;  Gaps   30;

Qy       8 DG-NEVIFTLWAPYQKSVKLKVLEKGLYEMER------DEKGYFTITLNNVKVRDRYKYV 60
           || : | | ||||  : | : | :   ::  |       : | : : |   :   ||:
Db     133 DGVSGVNFRLWAPNARRVSI-VGDFNYWDGRRHPMRFHSKSGVWELFLPKASLGQLYKFE 191

Qy      61 LDD-------------ASEI-PDPASRYQPEGVHGPSQIIQESKEFNNETFLKKEDLII 105
           | |             :|:: || ||  |   :    :: :  |: |       : : |
Db     192 LIDCHGNLRLKADPFAFSSQLRPDTAS--QVSALPNVVEMTEARKKANQ----GNQPISI 245

Qy     106 YEIHVGTF--TPEGTF-----EGVIRKLDYLKDLGITAIEIMPIAQFPGKRDWGYDGVYL 158
           ||:|:|::    |  |      :     : |:|::| | || :|:::||    |||  : |
Db     246 YEVHLGSWRRNLENNFWLDYDQIADELIPYVKEMGFTHIEFLPLSEFPFDGSWGYQPLGL 305

Qy     159 YAVQNSYGGPEGFRKLVDEAHKKGLGVILDVVYNHVGPEGNYMVKLGPYFSQKYKTP--- 215
           |:  : :| || ||:||  ||: |: |||| |  |   : : :|        ::: |
Db     306 YSPTSRFGSPEAFRRLVKRAHEAGINVILDWVPGHFPSDTHGLVAFDGTALYEHEDPREG 365

Qy     216 ----WG-LTFNFDDAESDEVRKFILENVEYWIKEYNVDGFRLDAVHAII---------- 259
               |  | :|:     :||: |:  |  ||:: : ||| |:||| ::|
Db     366 YHQDWNTLIYNYG---RNEVKNFLSSNALYWLERFGVDGIRVDAVASMIYRDYSRAEGEW 422

Qy     260 ---DTSPKHILEEIADVVH--------KYNRIVIAESDLNDPRVVNPKEKCGYNIDAQWV 308
               :  || |  : |            | |||   :   | :| |  |   : :|
Db     423 IPNQYGGRENLEAIEFLKHTNWKIHSEMAGAISIAEESTSFAGVTHPSENGGLGFNFKWN 482

Qy     309 DDFHHSIHAYLTGERQGYYTDFGNLDDIVKSY---KDVFVYDGKYSNFRRKTHGEPVGEL 365
             : :   ||:             || | : |   |  |    :||
Db     483 MGWMNDTLAYM------------KLDPIYRQYHHNKMTFGMVYQYSE------------ 517

Qy     366 DGCNFVVYIQNHDQVGNRGKGERIIKLVDRESYKIA--AALYLLS---PYIPMIFMGEEY 420
           |||: : :||:| : ||   : |:      | |   | |      |   ::||| |:
Db     518 ---NFVLPL-SHDEVVH-GKYSLLGKMPGDTWQKFANLRAYYGYMWGYPGKKLLFMGNEF 572

Qy     421 GEENPFYFFSDFSDSKLIQGVREGRKKENGQDTDPQDESTFNASKLSW-KIDEEI-FSFY 478
                               || ||   :|:                 | |  :|| |   ::
Db     573 A----------------QG-REWNYEES---------------LDWFLLDENIGGGWH 598

Qy     479 KILIKMRKELSIACDRR------VNVVNGENWLIIKGREYFSLYVFSKSSIEVKYSGTLL 532
           | ::|: |:|:    :        |   | :||::      |:  | : |   :    ::
Db     599 KGVLKLVKDLNQIYQKNRPLFELDNSPEGFDWLVVDDAAN-SVLAFERRSSNGE--RIIV 655

Qy     533 LSSNNSFPQHIEE------GKYE 549
           :|:    |:|         ||||
Db     656 VSNFTPVPRHNYRIGVNVAGKYE 678


RESULT  10
GLGB_STRAU
ID   GLGB_STRAU     STANDARD;      PRT;   764 AA.
AC   P52980;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   1,4-alpha-glucan branching enzyme (EC 2.4.1.18) (Glycogen branching
DE   enzyme).
GN   GLGB.
OS   Streptomyces aureofaciens.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=1894;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CCM 3239 / ATCC 10762;
RX   MEDLINE=94347823; PubMed=8068720;
RA   Homerova D., Kormanec J.;
RT   "Cloning of the putative glycogen branching enzyme gene, glgB, from
RT   Streptomyces aureofaciens.";
RL   Biochim. Biophys. Acta 1200:334-336(1994).
CC   -!- CATALYTIC ACTIVITY: Formation of 1,6-glucosidic linkages of
CC       glycogen.
CC   -!- PATHWAY: THIRD STEP IN GLYCOGEN BIOSYNTHESIS.
CC   -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC       KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
```

```
CC   ----------------------------------------------------------------------
DR   EMBL; L11647; AAA67437.1; -.
DR   InterPro; IPR000461; Alpha_amylase.
DR   InterPro; IPR001899; Gram_pos_anchor.
DR   InterPro; IPR004193; isoamylase_N.
DR   Pfam; PF00128; alpha-amylase; 1.
DR   Pfam; PF02922; isoamylase_N; 1.
KW   Glycogen biosynthesis; Transferase; Glycosyltransferase.
FT   ACT_SITE    440    440       BY SIMILARITY.
FT   ACT_SITE    493    493       BY SIMILARITY.
FT   ACT_SITE    561    561       BY SIMILARITY.
SQ   SEQUENCE   764 AA;  85325 MW;  6B45482E4A268ACF CRC64;


  Query Match             11.2%;  Score 335;  DB 1;  Length 764;
  Best Local Similarity   24.2%;  Pred. No. 1.5e-14;
  Matches  133;  Conservative   90;  Mismatches  197;  Indels  130;  Gaps   26;

Qy      14 FTLWAPYQKSVKL----KVLEKGLYEMER-DEKGYFTITLNNVKVRDRYKYVL---DDAS 65
           ||:|||    |::       :   | |      | : : |  |     ||| :   |
Db     178 FTVWAPNALGVRVTGDFSYWDAVAYPMRSLGASGVWELFLPGVAEGALYKYEITRPDGGR 237

Qy      66 EI-PDPASRYQPEGVHGPSQIIQESKEFNNETFLKKEDLI--------IYEIHVGTFTPE 116
           :  || :||        | :     |: :  :: :   :        :||:|: :: |
Db     238 TLRADPMARYAEVPPANASIVTASRYEWQDAEWMARRGALAPHQAPMSVYELHLASWRPG 297

Qy     117 GTFEGVIRKLD-YLKDLGITAIEIMPIAQFPGKRDWGYDGVYLYAVQNSYGGPEGFRKLV 175
           ::  :  :|  |:|:|| | :|:||:|: |    |||     ||  :  | |: || ||
Db     298 LSYRQLAEQLPAYVKELGFTHVELMPVAEHPFGGSWGYQVTGFYAPTSRMGTPDDFRFLV 357

Qy     176 DEAHKKGLGVILDVVYNHVGPEGNYMVKLGPYFSQKYKTP-------WGLTFNFDDAESD 228
           |  |: |:|||:| |  |   :   : :      ::: |       || |  ||
Db     358 DALHRAGIGVIVDWVPAHFPRDDWALAEFDGRPLYEHQDPRRAAHPDWG-TLEFDYGRK- 415

Qy     229 EVRKFILENVEYWIKEYNVDGFRLDAVHAII---------DTSPK-----------HILE 268
           ||| |:: |  || :|::||| | ||| :::          | ||              :|:
Db     416 EVRNFLVANAVYWCQEFHVDGLRADAVASMLYLDYSRDEGDWSPNAHGGREDLDAVALLQ 475

Qy     269 EI-ADVVHKYNRIV-IAESDLNDPRVVNPKEKCGYNIDAQWVDDFHHSIHAYLTGE---R 323
           |: | |  ::  :| |||        |  | :  |     :|   : |    |:: |   |
Db     476 EMNATVYRRFPGVVTIAEESTAWDGVTRPTDSGGLGFGLKWNMGWMHDTLRYVSKEPVHR 535

Qy     324 QGYYTD--FGNLDDIVKSYKDVFVYDGKYSNFRRKTHGEPVGELDGCNFVVYIQNHDQVG 381
           : :: |  ||    :| :: :                        |||: | :||:|
Db     536 KYHHHDMTFG----MVYAFSE-------------------------NFVLPI-SHDEVV 564

Qy     382 NRGKGERIIKL-----VDRESYKIAAALYLLSPYIPMIFMGEEY--GEE-----NPFYFF 429
           : ||   : |:       | :::        |   ::|||:|:  | |       | ::
Db     565 H-GKRSLVSKMPGDWWQQRATHRAYLGFMWAHPGKQLLFMGQEFAQGSEWSETYGPDWWV 623

Qy     430 SDFS-----DSKLIQGVREGRKKENGQDTDPQDESTFNASKLSWK--------------- 469
           | |     |     |||   :  |          |: ||   |:
Db     624 LDSSYPAAGDH---LGVRSLVRDLN---------RTYTASPALWERDSVPEGFAWVEADA 671

Qy     470 IDEEIFSFYK 479
           |: :|:| :
Db     672 ADDNVFAFLR 681


RESULT  11
GLGB_SYNY3
ID   GLGB_SYNY3     STANDARD;      PRT;   770 AA.
AC   P52981;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   1,4-alpha-glucan branching enzyme (EC 2.4.1.18) (Glycogen branching
DE   enzyme).
GN   GLGB OR SLL0158.
OS   Synechocystis sp. (strain PCC 6803).
OC   Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX   NCBI_TaxID=1148;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96127529; PubMed=8590279;
RA   Kaneko T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N.,
RA   Sugiura M., Tabata S.;
RT   "Sequence analysis of the genome of the unicellular cyanobacterium
RT   Synechocystis sp. strain PCC6803. I. Sequence features in the 1 Mb
RT   region from map positions 64% to 92% of the genome.";
RL   DNA Res. 2:153-166(1995).
CC   -!- CATALYTIC ACTIVITY: Formation of 1,6-glucosidic linkages of
CC       glycogen.
CC   -!- PATHWAY: THIRD STEP IN GLYCOGEN BIOSYNTHESIS.
CC   -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC       KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; D63999; BAA10073.1; -.
DR   InterPro; IPR000461; Alpha_amylase.
DR   InterPro; IPR004193; isoamylase_N.
DR   Pfam; PF00128; alpha-amylase; 1.
DR   Pfam; PF02922; isoamylase_N; 2.
KW   Glycogen biosynthesis; Transferase; Glycosyltransferase;
KW   Complete proteome.
FT   ACT_SITE    433    433       BY SIMILARITY.
FT   ACT_SITE    486    486       BY SIMILARITY.
FT   ACT_SITE    554    554       BY SIMILARITY.
SQ   SEQUENCE   770 AA;  89527 MW;  A435AFCA7703FA8A CRC64;


  Query Match             11.2%;  Score 333.5;  DB 1;  Length 770;
  Best Local Similarity   23.1%;  Pred. No. 1.9e-14;
  Matches  119;  Conservative   91;  Mismatches  176;  Indels  129;  Gaps   20;

Qy       7 IDG-NEVIFTLWAPYQKSVKL----KVLEKGLYEMERDEKGYFTITLNNVKVRDRYKYVL 61
           :||   | | :|||  ::| :       :  |::| :     : : :  : |   ||| :
Db     136 VDGVKGVYFAVWAPNARNVSILGDFNNWDGRLHQMRKRNNMVWELFIPELGVGTSYKYEI 195

Qy      62 DD----ASEIPDPASRYQPEGVHGPSQIIQ-ESKEFNNETFLK--------KEDLIIYEI 108
            :      |  ||   ||       | :   :  ::::| :|:          : : :||:
Db     196 KNWEGHIYEKTDPYGFYQEVRPKTASIVADLDGYQWHDEDWLEARRTSDPLSKPVSVYEL 255

Qy     109 HVG------------TFTPEGT---------------FEGVIRKLDYLKDLGITAIEIM 140
           |:|             |   ||                 :| | : : |:|:|| | ||::
Db     256 HLGSWLHTAYDEPVKTLHGEGVPVEVSEWNTGARFLTYYELVDKLIPYVKELGYTHIELL 315

Qy     141 PIAQFPGKRDWGYDGVYLYAVQNSYGGPEGFRKLVDEAHKKGLGVILDVVYNHVGPEGN- 199
           |||: |    |||      ||  : :| || |   ||: |  |:|||:| |  |   :|:
Db     316 PIAEHPFDGSWGYQVTGYYAPTSRFGSPEDFMYFVDQCHLNGIGVIIDWVPGHFPKDGHG 375

Qy     200 -------YMVKLGPYFSQKYKTPWG-LTFNFDDAESDEVRKFILENVEYWIKEYNVDGFR 251
                 :: : |     ::|  || | ||:     :||| |:: |  :|  :|::|| |
Db     376 LAFFDGTHLYEHGDPRKGEHK-EWGTLIFNYG---RNEVRNFLVANALFWFDKYHIDGMR 431

Qy     252 LDAVHAII-------------------DTSPKHILEEIADVVHKY--NRIVIAESDLND 289
           :||| :::                     :     | ::  ||: |    : |||   :
Db     432 VDAVASMLYLDYCREEGEWVANEYGGRENLEAADFLRQVNSVVYSYFPGILSIAEESTSW 491

Qy     290 PRVVNPKEKCGYNIDAQWVDDFHHSIHAYLTGERQGYYTDFGNLDDIVKSYKDVFVYDGK 349
           | |  |    |   : :|                     : | : |::    | |  |
Db     492 PMVSWPTYVGGLGFNLKW--------------------NMGWMHDML----DYFSMDPW 526

Qy     350 YSNFRRKT--------HGEPVGELDGCNFVVYIQNHDQVGNRGKGERIIKLVDRESYKIA 401
           :  | : :         | |        |::: : :||:| : ||   : |:   |  | |
Db     527 FRQFHQNSITFSMWYNHSE--------NYMLAL-SHDEVVH-GKSNMLGKMPGDEWQKYA 576

Qy     402 AALYLLS-----PYIPMIFMGEEYGEENPFYFFSD 431
```

```
              | :     |    :||  |:|: : :  : |
Db    577 NVRALFTYMFTHPGKKTMFMSMEFGQWSEWNVWGD 611


RESULT  12
GLGB_SYNP7
ID   GLGB_SYNP7     STANDARD;      PRT;   773 AA.
AC   P16954;
DT   01-AUG-1990 (Rel. 15, Created)
DT   01-AUG-1990 (Rel. 15, Last sequence update)
DT   01-FEB-1995 (Rel. 31, Last annotation update)
DE   1,4-alpha-glucan branching enzyme (EC 2.4.1.18) (Glycogen branching
DE   enzyme).
GN   GLGB.
OS   Synechococcus sp. (strain PCC 7942) (Anacystis nidulans R2).
OC   Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
OX   NCBI_TaxID=1140;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90323609; PubMed=2142668;
RA   Kiel J.A.K.W., Boels J.M., Beldman G., Venema G.;
RT   "Nucleotide sequence of the Synechococcus sp. PCC7942 branching
RT   enzyme gene (glgB): expression in Bacillus subtilis.";
RL   Gene 89:77-84(1990).
CC   -!- CATALYTIC ACTIVITY: Formation of 1,6-glucosidic linkages of
CC       glycogen.
CC   -!- PATHWAY: THIRD STEP IN GLYCOGEN BIOSYNTHESIS.
CC   -!- SUBUNIT: MONOMER.
CC   -!- MISCELLANEOUS: THE TEMPERATURE FOR OPTIMAL ACTIVITY IS
CC       APPROXIMATELY 35 DEGREES CELSIUS.
CC   -!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC       KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M31544; AAB39038.1; -.
DR   PIR; JQ0550; JQ0550.
DR   InterPro; IPR000461; Alpha_amylase.
DR   InterPro; IPR004193; isoamylase_N.
DR   Pfam; PF00128; alpha-amylase; 1.
DR   Pfam; PF02922; isoamylase_N; 1.
KW   Glycogen biosynthesis; Transferase; Glycosyltransferase.
FT   INIT_MET      0      0
FT   ACT_SITE    439    439       BY SIMILARITY.
FT   ACT_SITE    492    492       BY SIMILARITY.
FT   ACT_SITE    560    560       BY SIMILARITY.
SQ   SEQUENCE   773 AA;  89063 MW;  52BAA17CA337BF57 CRC64;


  Query Match             11.0%;  Score 326.5;  DB 1;  Length 773;
  Best Local Similarity   24.2%;  Pred. No. 5.3e-14;
  Matches  121;  Conservative   78;  Mismatches  175;  Indels  125;  Gaps   20;

Qy      12 VIFTLWAPYQKSVKL----KVLEKGLYEMERDEKGYFTITLNNVKVRDRYKYVLDDAS-- 65
           | | :|||  ::| :       :   ::| |   | : : :  : |   ||| : :
Db     147 VNFAVWAPSARNVSILGDFNSWDGRKHQMARRSNGIWELFIPELTVGAAYKYEIKNYDGH 206

Qy      66 --EIPDPASRYQPEGVHGPSQIIQESKEF--NNETFLK--------KEDLIIYEIHVGTF 113
             |  ||   :| |     : |: :   :   :  :|:        :: : :||:|:|::
Db     207 IYEKSDPYG-FQQEVRPKTASIVADLDRYTWGDADWLERRRHQEPLRQPISVYEVHLGSW 265

Qy     114 -------------------------TPEGTF----EGVIRKLDYLKDLGITAIEIMPIAQ 144
                                    |   |    |   | : |: ||| : ||::|||:
Db     266 MHASSDAIATDAQGKPLPPVPVADLKPGARFLTYRELADRLIPYVLDLGYSHIELLPIAE 325

Qy     145 FPGKRDWGYDGVYLYAVQNSYGGPEGFRKLVDEAHKKGLGVILDVVYNHVGPEGNYMVKL 204
           |    |||     ||  : || || |   ||  |: |:|||| |  |   :|: :

Db     326 HPFDGSWGYQVTGYYAATSRYGSPEDFMYFVDRCHQNGIGVILDWVPGHFPKDGHGLA-- 383

Qy     205 GPYFS----------QKYKTPWG-LTFNFDDAESDEVRKFILENVEYWIKEYNVDGFRL 252
             :|            |     || | ||:      ||| |:  |  :|  :|::|| |:
Db     384 --FFDGTHLYEHADSRQGEHREWGTLVFNYG---RHEVRNFLAANALFWFDKYHIDGIRV 438

Qy     253 DAVHAI--IDTSPK------------HILEEIADVVHKYNRIV---------IAESDLN 288
           ||| ::  :| : |              :|  | || : : | ::         |||   :
Db     439 DAVASMLYLDYNRKEGEWIPNEYGGRENI--EAADFLRQVNHLIFSYFPGALSIAEESTS 496

Qy     289 DPRVVNPKEKCGYNIDAQWVDDFHHSIHAYLTGERQGYYTDFGNLDDIVKSYKDVFVYDG 348
           | |  |    |   : :|                      : | : |::    | |  |
Db     497 WPMVSWPTYVGGLGFNLKW--------------------NMGWMHDML----DYFSMDP 531

Qy     349 KYSNFRRKTHGEPVGELDGCNFVVYIQNHDQVGNRGKGERIIKLVDRESYKIAAALYLLS 408
           :  | :      :      ||:: : :||:| : ||   | |:   |  | |    ||
Db     532 WFRQFHQNNVTFSIWYAFSENFMLAL-SHDEVVH-GKSNLIGKMPGDEWQKFANLRCLLG 589

Qy     409 -----PYIPMIFMGEEYGE 422
                |    :||| |:|:
Db     590 YMFTHPGKKTLFMGMEFGQ 608


RESULT  13
GLGB_MYCTU
ID   GLGB_MYCTU     STANDARD;      PRT;   731 AA.
AC   Q10625;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Probable 1,4-alpha-glucan branching enzyme (EC 2.4.1.18) (Glycogen
DE   branching enzyme).
GN   GLGB OR RV1326C OR MT1368 OR MTCY130.11C.
OS   Mycobacterium tuberculosis.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX   NCBI_TaxID=1773;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=H37RV;
RX   MEDLINE=98295987; PubMed=9634230;
RA   Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA   Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekaia F.,
RA   Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA   Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA   Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA   Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA   Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA   Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT   "Deciphering the biology of Mycobacterium tuberculosis from the
RT   complete genome sequence.";
RL   Nature 393:537-544(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CDC 1551 / Oshkosh;
RA   Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA   Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA   Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA   Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA   Bishai W.;
RT   "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT   laboratory strains.";
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- CATALYTIC ACTIVITY: Formation of 1,6-glucosidic linkages of
CC       glycogen.
CC   -!- PATHWAY: THIRD STEP IN GLYCOGEN BIOSYNTHESIS.
CC   -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC       KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
```

```
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z73902; CAA98090.1; -.
DR   EMBL; AE007010; AAK45632.1; -.
DR   TIGR; MT1368; -.
DR   TubercuList; Rv1326c; -.
DR   InterPro; IPR000461; Alpha_amylase.
DR   InterPro; IPR004193; isoamylase_N.
DR   Pfam; PF00128; alpha-amylase; 1.
DR   Pfam; PF02922; isoamylase_N; 1.
KW   Glycogen biosynthesis; Transferase; Glycosyltransferase;
KW   Complete proteome.
FT   ACT_SITE    411    411       BY SIMILARITY.
FT   ACT_SITE    464    464       BY SIMILARITY.
FT   ACT_SITE    532    532       BY SIMILARITY.
FT   CONFLICT    214    214       P -> A (IN REF. 2).
FT   CONFLICT    223    223       P -> Q (IN REF. 2).
SQ   SEQUENCE   731 AA;  81729 MW;  EE2BFEF765352617 CRC64;


  Query Match           10.9%;  Score 326;  DB 1;  Length 731;
  Best Local Similarity  22.8%;  Pred. No. 5.4e-14;
  Matches  117;  Conservative   76;  Mismatches  186;  Indels  134;  Gaps   18;

Qy      12 VIFTLWAPYQKSVKL-------------KVL-EKGLYEMERDE---KGYFTITLNNVK- 52
           | | :|||  | | |             :||   |::|:   :    | :   ::
Db     147 VSFAVWAPNAKGVSLIGEFNGWNGHEAPMRVLGPSGVWELFWPDFPCDGLYKFRVHGADG 206

Qy      53 -VRDRYKYVLDDASEIPDPASRYQPEGVHGPSQIIQESKEFNNETFLKKEDLIIYEIHVG 111
            | ||           |  |||                :   |      | :  ||:|:|
Db     207 VVTDRADPFAFGTEVPPQTASRVTSSDYTWGDDDWMAGRALRNPV---NEAMSTYEVHLG 263

Qy     112 TFTPEGTFEGVIRKL-DYLKDLGITAIEIMPIAQFPGKRDWGYDGVYLYAVQNSYGGPEG 170
           :: |  ::  : |:| ||: | | | :|::|:|: |    |||     ||  : :| |:
Db     264 SWRPGLSYRQLARELTDYIVDQGFTHVELLPVAEHPFAGSWGYQVTSYYAPTSRFGTPDD 323

Qy     171 FRKLVDEAHKKGLGVILDVV-------------------YNHVGPEGNYMVKLGPYFSQK 211
           || |||  |: |:|||:| |                   | |  |:    :  | |
Db     324 FRALVDALHQAGIGVIVDWVPAHFPKDAWALGRFDGTPLYEHSDPKRGEQLDWGTY---- 379

Qy     212 YKTPWGLTFNFDDAESDEVRKFILENVEYWIKEYNVDGFRLDAVHAII------------ 259
                 |:|      ||| |:: |  ||::|:::|| |:||| :::
Db     380 -------VFDFG---RPEVRNFLVANALYWLQEFHIDGLRVDAVASMLYLDYSRPEGGWT 429

Qy     260 --------DTSPKHILEEIADVVHKY--NRIVIAESDLNDPRVVNPKEKCGYNIDAQWVD 309
                   :      |:|:    ||     : |||       |  |    |     :|
Db     430 PNVHGGRENLEAVQFLQEMNATAHKVAPGIVTIAEESTPWSGVTRPTNIGGLGFSMKWNM 489

Qy     310 DFHHSIHAYLTGERQGYYTDFGNLDDIVKSYKD-------VFVYDGKYSNFRRKTHGEPV 362
            : |                |: : | : :||        :: :   |
Db     490 GWMHDT------------LDYVSRDPVYRSYHHHEMTFSMLYAFSENY------------ 525

Qy     363 GELDGCNFVVYIQNHDQVGNRGKGERIIKLVDRESYKIAAALYLLS-----PYIPMIFMG 417
                 |   :||:| : |||    ::      | |    ||:     |   ::|||
Db     526 ---------VLPLSHDEVVH-GKGTLWGRMPGNNHVKAAGLRSLLAYQWAHPGKQLLFMG 575

Qy     418 EEYGEENPFYFFSDFSDSKLIQGVREGRKKENG 450
           :|:|:       :::|: :   |:   :  |||
Db     576 QEFGQR------AEWSEQR---GLDWFQLDENG 599


RESULT  14
GLGB_BUTFI
ID   GLGB_BUTFI     STANDARD;      PRT;   639 AA.
AC   P30539;
DT   01-APR-1993 (Rel. 25, Created)
DT   01-APR-1993 (Rel. 25, Last sequence update)
DT   01-FEB-1995 (Rel. 31, Last annotation update)
DE   1,4-alpha-glucan branching enzyme (EC 2.4.1.18) (Glycogen branching
DE   enzyme).
GN   GLGB.
OS   Butyrivibrio fibrisolvens.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC   Butyrivibrio.
OX   NCBI_TaxID=831;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 1-12.
RC   STRAIN=H17C;
RX   MEDLINE=92041554; PubMed=1938880;
RA   Rumbak E., Rawlings D.E., Lindsey G.G., Woods D.R.;
RT   "Characterization of the Butyrivibrio fibrisolvens glgB gene, which
RT   encodes a glycogen-branching enzyme with starch-clearing activity.";
RL   J. Bacteriol. 173:6732-6741(1991).
CC   -!- FUNCTION: TRANSFERS CHAINS OF 5 TO 10 (OPTIMUM, 7) GLUCOSE UNITS,
CC       USING AMYLOSE AND AMYLOPECTIN AS SUBSTRATES, TO PRODUCE A HIGHLY
CC       BRANCHED POLYMER.
CC   -!- CATALYTIC ACTIVITY: Formation of 1,6-glucosidic linkages of
CC       glycogen.
CC   -!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC       KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M64980; AAA23007.1; -.
DR   PIR; B41328; B41328.
DR   InterPro; IPR000461; Alpha_amylase.
DR   InterPro; IPR004193; isoamylase_N.
DR   Pfam; PF00128; alpha-amylase; 1.
DR   Pfam; PF02922; isoamylase_N; 1.
KW   Glycogen biosynthesis; Transferase; Glycosyltransferase.
FT   ACT_SITE    313    313       BY SIMILARITY.
FT   ACT_SITE    366    366       BY SIMILARITY.
FT   ACT_SITE    434    434       BY SIMILARITY.
SQ   SEQUENCE   639 AA;  73875 MW;  DC93B72A5F28C2AF CRC64;


  Query Match           10.8%;  Score 321;  DB 1;  Length 639;
  Best Local Similarity  25.1%;  Pred. No. 9.5e-14;
  Matches  124;  Conservative   81;  Mismatches  153;  Indels  136;  Gaps   27;

Qy      14 FTLWAPYQKSVKL----KVLEKGLYEMERDEKG-YFTITLNNVKVRDRYKYVL--DDASE 66
           | :|||    | :        ::  ::|:| : |  :|: :  | :   ||:::   |  :
Db      42 FAVWAPNAADVHVVGDFNGWDENAHQMKRSKTGNIWTLFIPGVAIGALYKFLITAQDGRK 101

Qy      67 I--PDPASRY---QPEGVHGPSQII----QESKEFNNETFLKKED-----LIIYEIHVGT 112
           :   || : |   :|       : :      :|| : :   || :|     : ||| |:|:
Db     102 LYKADPYANYAELRPGNASRTTDLSGFKWSDSKWYES---LKGKDMNRQPIAIYECHIGS 158

Qy     113 FT--PEGTFEGVI-------RKLDYLKDLGITAIEIMPIAQFPGKRDWGYDGVYLYAVQN 163
           :   |:|| :|          | ::|||::  | ||:: ||: |    |||     ||
Db     159 WMKHPDGTEDGFYTYRQFADRIVEYLKEMKYTHIELIGIAEHPFDGSWGYQVTGYYAPTA 218

Qy     164 SYGGPEGFRKLVDEAHKKGLGVILDVVYNHVGP--------EGNYMVKLGPYFSQKYKTP 215
           || |  |  |::: || |:|||| |  |  |         :|  :           |: |
Db     219 RYGEPTDFMYLINQLHKHGIGVILDWVPAHFCPDEFGLACFDGTCI----------YEDP 268

Qy     216 ---------WGL-TFNFDDAESDEVRKFILENVEYWIKEYNVDGFRLDAVHAI--IDTSP 263
                     ||   ||       ||: |:: |  |||:::::|| |:||| ::  :|
Db     269 DPRKGEHPDWGTKIFNL---AKPEVKNFLIANALYWIRKFHIDGLRVDAVASMLYLDYGK 325

Qy     264 K------------HILEEIADVVHKYNRIV---------IAESDLNDPRVVNPKEKCGYN 302
           |              |: |    | :| :|          |||     |:|  | |: |
Db     326 KDGQWVPNKYGDNKNLDAIEFFKH-FNSVVRGTYPNILTIAEESTAWPKVTAPPEEDGLG 384

Qy     303 IDAQWVDDFHHSIHAYLTGE---RQG--YYTDFGNLDDIVKSYKDVFVYDGKYSNFRRKT 357
             :|   : |    |:  :   |||  |   |         ||
Db     385 FAFKWNMGWMHDFCEYMKLDPYFRQGAHYMMTF------AMSYN---------------- 422
```

```
Qy    358 HGEPVGELDGCNFVVYIQNHDQVGNRGKGERIIKLVDRESYKIAAALYLLSPYIPM---- 413
                |  |::: : :||:| :  |   : |:     ||:     |   |  |
Db    423 --------DSENYILPL-SHDEVVHL-KCSMVEKM---PGYKVDKYANLRVGYTYMFGHS 469

Qy    414 ----IFMGEEYGEE 423
              :|||:::|:|
Db    470 GKKLLFMGQDFGQE 483


RESULT 15
GLGX_HAEIN
ID   GLGX_HAEIN     STANDARD;      PRT;   659 AA.
AC   P45178;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Glycogen operon protein glgX (EC 3.2.1.-).
GN   GLGX OR HI1358.
OS   Haemophilus influenzae.
OC   Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC   Haemophilus.
OX   NCBI_TaxID=727;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=RD / KW20 / ATCC 51907;
RX   MEDLINE=95350630; PubMed=7542800;
RA   Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA   Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA   McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA   Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA   Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA   Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA   Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA   Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA   Venter J.C.;
RT   "Whole-genome random sequencing and assembly of Haemophilus
RT   influenzae Rd.";
RL   Science 269:496-512(1995).
CC   -!- FUNCTION: THIS PROTEIN MAY BE PART OF A GLYCOGEN BIOSYNTHETIC/
CC       CATABOLIC OPERON BUT IS NOT REQUIRED FOR GLYCOGEN SYNTHESIS
CC       (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC       KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U32815; AAC23005.1; -.
DR   HSSP; P10342; 1BF2.
DR   TIGR; HI1358; -.
DR   InterPro; IPR000461; Alpha_amylase.
DR   InterPro; IPR004193; isoamylase_N.
DR   Pfam; PF00128; alpha-amylase; 1.
DR   Pfam; PF02922; isoamylase_N; 1.
KW   Glycogen biosynthesis; Hydrolase; Glycosidase; Complete proteome.
FT   ACT_SITE    333    333       BY SIMILARITY.
FT   ACT_SITE    440    440       BY SIMILARITY.
SQ   SEQUENCE   659 AA;  75290 MW;  E2B71F610E8B1CAC CRC64;


  Query Match            10.7%;  Score 318;  DB 1;  Length 659;
  Best Local Similarity  23.1%;  Pred. No. 1.6e-13;
  Matches  137;  Conservative   98;  Mismatches  173;  Indels  186;  Gaps   34;

Qy      3 FAYKIDG-----NEVIFTLWAPYQKSVKLKVLEKGLYEMERDE-KGYFTITLNNVKVRDR 56
          :|::| |      ::|     || |:|       |  ::  :| : :| :: |    ||
Db     74 YAFRIHGEFANPQKLILD---PYAKAV------NGKPDLSSEESRSWFLLSDN----RDN 120

Qy     57 YKYVLDDASEIPDPASRYQPEGVHGPSQIIQESKEFNNETFLKK--EDLIIYEIHVGTFT 114
                      :   |  |     :| |  :: |:|       : |:||:||  |:
Db    121 ---------------AHLAPRAV-----VISEEFDWENDTSPNTPWAETIVYELHVKGFS 160

Qy    115 P---------EGTFEGVIR--KLDYLKDLGITAIEIMPI--------AQFPGKRD-WGYD 154
                    ||: |:     | |||:||:|||::|:        |  | :: |||:
Db    161 QLNEKIPAALRGTYTGLAHPVNLAYLKELGVTAVELLPVNFHINEPHLQARGLQNYWGYN 220

Qy    155 GVYLYAVQNSYGGPEG----FRKLVDEAHKKGLGVILDVVYNH-VGPEGNYMVKLGPYFS 209
           : ::||:  |          |: :|   || |: |||||:||    |  |     | ||
Db    221 PLAMFAVEPKYAATNNPLAEFKTMVKAFHKAGIEVILDVVFNHSAESEQTY-----PTFS 275

Qy    210 QK--------YKTPWGLTFNFD------DAESDEVRKFILENVEYWIKEYNVDGFRLDAV 255
          |:        ::   |   |:       :  ||  ||:::: : ||::: ::|||| |
Db    276 QRGIDDQTYYWRNDQGRYINWTGCGNMLNLSSDVGRKWVVDCLRYWVEQCHIDGFRFDLA 335

Qy    256 HAIIDTSPKHILEEIADVVHKYNRIVIAESDL-NDPRVVNPK--------EKCGYNID-- 304
            :  :|          :|     :|: |:| : | |           || :
Db    336 TVLGRDTP------------DFNSSAQLFTDIKNEPSLQNIKLIAEPWDIGHYGYQVGNF 383

Qy    305 ----AQWVDDFHHSI---HAYLTGERQGYYTDFGNLDDIVK------------------- 338
              |:| | |   :     : :||   :   |    |: |
Db    384 PSYFAEWNDRFRDDLCRFWLWKSGEIGAFAERFAGSSDLFKKNDRLPHTTLNFITAHDGF 443

Qy    339 SYKDVFVYDGKYSNFRRKTHGEPVGELDGCNFVVYIQNHDQVGNRGKGERIIK-----LV 393
          : ||:  |: |::    :|:||     || |   |  ||   |  |  | : :     :
Db    444 TLKDLVSYNQKHN----ETNGEE--NRDGRN-ENYSYNH---GVEGSTESLSEPQKSAVE 493

Qy    394 DRESYKIAAAL--YLLSPYIPMIFMGEEYGEENPFYFFSDFSDSKLIQGVREGRKKENGQ 451
          :  ::  :  |   ||:   ||:  |:|:|                 : |      |
Db    494 NNRTFAQSGLLMSLLLANGTPMLLAGDEFG-----------------NTQYGNNNAYCQ 535

Qy    452 DTDPQDESTFNASKLSW-KIDEEIFSFYKILIKMRKELSIACDRRVNVVNGENW 504
          |    :| |:    | |   :||:|   |  | :||:        :  :| : |
Db    536 D----NEITW----LKWANFNEELFELTKQTIALRKQ--------IGSLNKDQW 573


Search completed: July 15, 2002, 12:24:32
Job time: 384 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: July 15, 2002, 12:17:38 ; Search time 55.94 Seconds
(without alignments)
1728.712 Million cell updates/sec

Title: US-09-298-924-6
Perfect score: 2978
Sequence: 1 MTFAYKIDGNEVIFTLWAPY.....PQHIEEGKYEFDKGFALYKL 559

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_19:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2978	100.0	559	1	Q55088	Q55088 sulfolobus
2	2837	95.3	559	1	Q9UWN9	Q9uwn9 sulfolobus
3	2467	82.8	561	17	P95867	P95867 sulfolobus
4	1748.5	58.7	556	1	Q53641	Q53641 sulfolobus
5	1269	42.6	359	17	Q973H4	Q973h4 sulfolobus
6	1033.5	34.7	581	2	Q9ADI5	Q9adi5 streptomyce
7	905.5	30.4	600	16	Q9RX51	Q9rx51 deinococcus
8	881.5	29.6	583	16	Q9I1V1	Q9i1v1 pseudomonas
9	876.5	29.4	601	16	Q92U63	Q92u63 rhizobium m
10	706	23.7	217	17	Q973H5	Q973h5 sulfolobus
11	524.5	17.6	552	2	Q44528	Q44528 anabaena va
12	399	13.4	652	2	Q59242	Q59242 bacillus st
13	397	13.3	440	2	Q93Q35	Q93q35 myxococcus
14	395.5	13.3	1142	2	Q93UZ9	Q93uz9 bacillus sp
15	381.5	12.8	825	2	Q59319	Q59319 caldocellum
16	365	12.3	1280	16	Q97SQ7	Q97sq7 streptococc
17	364	12.2	1287	2	Q9F930	Q9f930 streptococc
18	361	12.1	630	16	O66936	O66936 aquifex aeo
19	349.5	11.7	783	10	O04196	O04196 arabidopsis
20	348.5	11.7	666	16	O84046	O84046 chlamydia t
21	346.5	11.6	713	1	O05152	O05152 sulfolobus
22	345.5	11.6	422	2	Q59243	Q59243 bacillus st
23	345.5	11.6	664	16	Q9Z8F5	Q9z8f5 chlamydia p
24	345.5	11.6	666	16	Q9PKZ6	Q9pkz6 chlamydia m
25	345.5	11.6	707	16	P73608	P73608 synechocyst
26	344.5	11.6	741	2	Q9KY06	Q9ky06 streptomyce
27	344.5	11.6	741	2	Q59832	Q59832 streptomyce
28	343	11.5	718	17	P95868	P95868 sulfolobus
29	341	11.5	737	16	Q985P4	Q985p4 rhizobium l
30	339.5	11.4	789	10	O22637	O22637 zea mays (m
31	336	11.3	702	10	Q9M0S5	Q9m0s5 arabidopsis
32	336	11.3	716	17	Q973H3	Q973h3 sulfolobus
33	335	11.2	1072	16	Q9K6N1	Q9k6n1 bacillus ha
34	334.5	11.2	717	16	Q9K7U5	Q9k7u5 bacillus ha
35	334.5	11.2	736	16	Q92M14	Q92m14 rhizobium m
36	333.5	11.2	720	16	Q9RXP5	Q9rxp5 deinococcus
37	331.5	11.1	1938	2	P70983	P70983 bacillus sp
38	330.5	11.1	562	2	Q9L872	Q91872 bacillus sp
39	330.5	11.1	562	16	Q9KTJ1	Q9ktj1 vibrio chol
40	329.5	11.1	668	2	P71095	P71095 bacteroides
41	329.5	11.1	818	10	Q41742	Q41742 zea mays (m
42	329.5	11.1	1165	16	Q99XX8	Q99xx8 streptococc
43	329	11.0	741	3	Q9P5P3	Q9p5p3 neurospora
44	327	11.0	733	10	O80403	O80403 oryza sativ
45	323.5	10.9	718	16	O34587	O34587 bacillus su

ALIGNMENTS

RESULT 1
Q55088
ID Q55088 PRELIMINARY; PRT; 559 AA.
AC Q55088;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE ALPHA-AMYLASE.
OS Sulfolobus solfataricus.
OC Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.
OX NCBI_TaxID=2287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KM1;
RA Kato M., Kettoku M., Miura Y., Komeda T., Konishi Y., Shindo K.,
RA Kobayashi K., Iwamatsu A.;
RT "The gene analysis of the new amylases from the hyper thermophilic
RT archae Sulfolobus.";
RL Submitted (SEP-1995) to the EMBL/GenBank/DDBJ databases.
DR EMBL; D64130; BAA11010.1; -.
DR InterPro; IPR000461; Alpha_amylase.
DR InterPro; IPR004193; isoamylase_N.
DR Pfam; PF00128; alpha-amylase; 1.
DR Pfam; PF02922; isoamylase_N; 1.
SQ SEQUENCE 559 AA; 64790 MW; 79FBE23A7CD38B4E CRC64;

Query Match 100.0%; Score 2978; DB 1; Length 559;
Best Local Similarity 100.0%; Pred. No. 1.3e-180;
Matches 559; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 MTFAYKIDGNEVIFTLWAPYQKSVKLKVLEKGLYEMERDEKGYFTITLNNVKVRDRYKYV 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MTFAYKIDGNEVIFTLWAPYQKSVKLKVLEKGLYEMERDEKGYFTITLNNVKVRDRYKYV 60

Qy      61 LDDASEIPDPASRYQPEGVHGPSQIIQESKEFNNETFLKKEDLIIYEIHVGTFTPEGTFE 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 LDDASEIPDPASRYQPEGVHGPSQIIQESKEFNNETFLKKEDLIIYEIHVGTFTPEGTFE 120
```

```
Qy    121 GVIRKLDYLKDLGITAIEIMPIAQFPGKRDWGYDGVYLYAVQNSYGGPEGFRKLVDEAHK 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GVIRKLDYLKDLGITAIEIMPIAQFPGKRDWGYDGVYLYAVQNSYGGPEGFRKLVDEAHK 180

Qy    181 KGLGVILDVVYNHVGPEGNYMVKLGPYFSQKYKTPWGLTFNFDDAESDEVRKFILENVEY 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 KGLGVILDVVYNHVGPEGNYMVKLGPYFSQKYKTPWGLTFNFDDAESDEVRKFILENVEY 240

Qy    241 WIKEYNVDGFRLDAVHAIIDTSPKHILEEIADVVHKYNRIVIAESDLNDPRVVNPKEKCG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 WIKEYNVDGFRLDAVHAIIDTSPKHILEEIADVVHKYNRIVIAESDLNDPRVVNPKEKCG 300

Qy    301 YNIDAQWVDDFHHSIHAYLTGERQGYYTDFGNLDDIVKSYKDVFVYDGKYSNFRRKTHGE 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 YNIDAQWVDDFHHSIHAYLTGERQGYYTDFGNLDDIVKSYKDVFVYDGKYSNFRRKTHGE 360

Qy    361 PVGELDGCNFVVYIQNHDQVGNRGKGERIIKLVDRESYKIAAALYLLSPYIPMIFMGEEY 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 PVGELDGCNFVVYIQNHDQVGNRGKGERIIKLVDRESYKIAAALYLLSPYIPMIFMGEEY 420

Qy    421 GEENPFYFFSDFSDSKLIQGVREGRKKENGQDTDPQDESTFNASKLSWKIDEEIFSFYKI 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GEENPFYFFSDFSDSKLIQGVREGRKKENGQDTDPQDESTFNASKLSWKIDEEIFSFYKI 480

Qy    481 LIKMRKELSIACDRRVNVVNGENWLIIKGREYFSLYVFSKSSIEVKYSGTLLLSSNNSFP 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 LIKMRKELSIACDRRVNVVNGENWLIIKGREYFSLYVFSKSSIEVKYSGTLLLSSNNSFP 540

Qy    541 QHIEEGKYEFDKGFALYKL 559
          |||||||||||||||||||
Db    541 QHIEEGKYEFDKGFALYKL 559
```

RESULT 2
Q9UWN9
ID Q9UWN9 PRELIMINARY; PRT; 559 AA.
AC Q9UWN9;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE MALTOOLIGOSYL TREHALOSE TREHALOHYDROLASE.
GN TREZ.
OS Sulfolobus shibatae.
OC Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.
OX NCBI_TaxID=2286;
RN [1]
RP SEQUENCE FROM N.A.
RA Weiting Y., Wei C., Hui W., Li L., Cheng J.;
RT "Maltooligosyl trehalose trehalohydrolase from Sulfolobus shibatae.";
RL Submitted (NOV-1999) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF201335; AAF17553.1; -.
DR InterPro; IPR000461; Alpha_amylase.
DR InterPro; IPR004193; isoamylase_N.
DR Pfam; PF00128; alpha-amylase; 1.
DR Pfam; PF02922; isoamylase_N; 1.
KW Hydrolase.
SQ SEQUENCE 559 AA; 64619 MW; 8C0F0F38A5AE2846 CRC64;

Query Match 95.3%; Score 2837; DB 1; Length 559;
Best Local Similarity 95.2%; Pred. No. 1.1e-171;
Matches 532; Conservative 11; Mismatches 16; Indels 0; Gaps 0;

```
Qy      1 MTFAYKIDGNEVIFTLWAPYQKSVKLKVLEKGLYEMERDEKGYFTITLNNVKVRDRYKYV 60
          ||| |||||||||||||||||||||||||||||||||:||||||||||||| ||||||
Db      1 MTFGYKIDGNEVIFTLWAPYQKSVKLKVLEKGLYEMERDDKGYFTITLNNVKVGDRYKYV 60

Qy     61 LDDASEIPDPASRYQPEGVHGPSQIIQESKEFNNETFLKKEDLIIYEIHVGTFTPEGTFE 120
          |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db     61 LDDASEIPDPASRYQPEGVHGPSQIIQESKEFNNETFLKKEDLIIYEIHVGTFAPEGTFE 120

Qy    121 GVIRKLDYLKDLGITAIEIMPIAQFPGKRDWGYDGVYLYAVQNSYGGPEGFRKLVDEAHK 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    121 GVIRKLDYLKDLGITAIEIMPIAQFPGKRDWGYDGVYLYAVQNSYGGPEGFRKLVDEAHM 180

Qy    181 KGLGVILDVVYNHVGPEGNYMVKLGPYFSQKYKTPWGLTFNFDDAESDEVRKFILENVEY 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 KGLGVILDVVYNHVGPEGNYMVKLGPYFSQKYKTPWGLTFNFDDAESDEVRKFILENVEY 240

Qy    241 WIKEYNVDGFRLDAVHAIIDTSPKHILEEIADVVHKYNRIVIAESDLNDPRVVNPKEKCG 300
          ||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||
Db    241 WIKEYNVDGFRLDAVHSIIDTSPKHILEEIADVVHKYNRIVIAESDLNDPRVVNPKEKCG 300

Qy    301 YNIDAQWVDDFHHSIHAYLTGERQGYYTDFGNLDDIVKSYKDVFVYDGKYSNFRRKTHGE 360
          |||||||||||||||||||||||||| ||||||:|||||||||||||| |||||||:
Db    301 YNIDAQWVDDFHHSIHAYLTGERQGYYLDFGNLDNIVKSYKDVFVYDGKYSKFRRKTHGK 360

Qy    361 PVGELDGCNFVVYIQNHDQVGNRGKGERIIKLVDRESYKIAAALYLLSPYIPMIFMGEEY 420
          |||||||||||||||||||||||||||||||:|||| |||||:|| | ||||||||
Db    361 PVGELDGCNFVVYIQNHDQVGNRGKGERIIKLVDKESYKNAAALYILSSIISMIFMGEEY 420

Qy    421 GEENPFYFFSDFSDSKLIQGVREGRKKENGQDTDPQDESTFNASKLSWKIDEEIFSFYKI 480
          ||: |||||||||:|||||:||||||||||||||||||||||||||||||||||||||||||
Db    421 GEQKPFYFFSDFSDTKLIQGLREGRKKENGQDTDPQDESTFNASKLSWKIDEEIFSFYKI 480

Qy    481 LIKMRKELSIACDRRVNVVNGENWLIIKGREYFSLYVFSKSSIEVKYSGTLLLSSNNSFP 540
          ||||| |||||||||| | |||||||||||||| || |||:||||||||||||||||||||
Db    481 LIKMRNELSIACDRRVYVENGENWLIIKGREYLSLCVFSRSSIEVKYSGTLLLSSNNSFP 540

Qy    541 QHIEEGKYEFDKGFALYKL 559
          |||||||:|||||||||||
Db    541 QHIEEGQYEFDKGFALYKL 559
```

RESULT 3
P95867
ID P95867 PRELIMINARY; PRT; 561 AA.
AC P95867;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE ALPHA-AMYLASE PRECURSOR (MALTO-OLIGOSYLTREHALOSE TREHALOHYDROLASE)
DE (TREZ) (EC 3.2.1.141).
GN TREZ.
OS Sulfolobus solfataricus.
OC Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.
OX NCBI_TaxID=2287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35092 / DSM 1617 / P2;
RX MEDLINE=21332296; PubMed=11427726;
RA She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
RA Awayez M.J., Chan-Weiher C.C.-Y., Clausen I.G., Curtis B.A.,
RA De Moors A., Erauso G., Fletcher C., Gordon P.M.K.,
RA Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
RA Garrett R.A., Ragan M.A., Sensen C.W., Van der Oost J.;
RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
DR EMBL; Y08256; CAA69503.1; -.
DR EMBL; AE006815; AAK42272.1; -.
DR InterPro; IPR000461; Alpha_amylase.
DR InterPro; IPR004193; isoamylase_N.
DR Pfam; PF00128; alpha-amylase; 1.
DR Pfam; PF02922; isoamylase_N; 1.
KW Hydrolase; Glycosidase; Complete proteome.
SQ SEQUENCE 561 AA; 64370 MW; B00E403020F6B242 CRC64;

Query Match 82.8%; Score 2467; DB 17; Length 561;
Best Local Similarity 79.7%; Pred. No. 3e-148;
Matches 447; Conservative 63; Mismatches 49; Indels 2; Gaps 1;

```
Qy      1 MTFAYKIDGNEVIFTLWAPYQKSVKLKVLEKGLYEMERDEKGYFTITLNNVKVRDRYKYV 60
          ||| ||:| : | | |||||: ||||:| :|:|||||:||||||||:||:| |||||:
```

```
Db      1 MTFGYKLDEDGVTFNLWAPYQRKVKLKILNRGIYEMERDDKGYFTITLDNVRVGDRYKYI 60

Qy     61 LDDASEIPDPASRYQPEGVHGPSQIIQESKEFNNETF--LKKEDLIIYEIHVGTFTPEGT 118
          ||| ||:|||||||||||||| |:|| |:::| :|:|||:|||:|:|||| |||
Db     61 LDDNSEVPDPASRYQPEGVHGYSEIISPDFEWDDENSVKVKREDLVIYELHIGTFTSEGT 120

Qy    119 FEGVIRKLDYLKDLGITAIEIMPIAQFPGKRDWGYDGVYLYAVQNSYGGPEGFRKLVDEA 178
          ||||||:||:|||:||:|||||||||||||:|||||||||||||||| ||||||:||
Db    121 FEGVIKKLNYLKELGVTAIEIMPIAQFPGKKDWGYDGVYLYAVQNSYGGPSGFRKLVNEA 180

Qy    179 HKKGLGVILDVVYNHVGPEGNYMVKLGPYFSQKYKTPWGLTFNFDDAESDEVRKFILENV 238
          || || |||||||||||||||||||||||:|||||||||||||| ||||||||||
Db    181 HKLGLAVILDVVYNHVGPEGNYMVKLGPYFSEKYKTPWGLTFNFDDAGSDEVRKFILENV 240

Qy    239 EYWIKEYNVDGFRLDAVHAIIDTSPKHILEEIADVVHKYNRIVIAESDLNDPRVVNPKEK 298
          |||| |::|||||||||||||| |||||||:||||||||::|||||||||||||||||||
Db    241 EYWINEFHVDGFRLDAVHAIIDNSPKHILEDIADVVHKYDKIVIAESDLNDPRVVNPKEK 300

Qy    299 CGYNIDAQWVDDFHHSIHAYLTGERQGYYTDFGNLDDIVKSYKDVFVYDGKYSNFRRKTH 358
          |||||||||||||||:|||:|||||||||:|||:: ||||||||||:||||||||||||||
Db    301 CGYNIDAQWVDDFHHAIHAFLTGERQGYYSDFGSIGDIVKSYKDVFIYDGKYSNFRRKTH 360

Qy    359 GEPVGELDGCNFVVYIQNHDQVGNRGKGERIIKLVDRESYKIAAALYLLSPYIPMIFMGE 418
          |: ||:|||| |||||||||||||| |||:|||||:|||||||||:||||||||||||
Db    361 GKSVGDLDGCKFVVYIQNHDQVGNRGGGERLIKLVDKESYKIAAALYILSPYIPMIFMGE 420

Qy    419 EYGEENPFYFFSDFSDSKLIQGVREGRKKENGQDTDPQDESTFNASKLSWKIDEEIFSFY 478
          |||||||||:|||||| |||||||||::||||:|||| : ||| |||||||:::| |||
Db    421 EYGEENPFYYFSDFSDPKLIQGVREGRRRENGQETDPQSDCTFNDSKLSWKINDDILSFY 480

Qy    479 KILIKMRKELSIACDRRVNVVNGENWLIIKGREYFSLYVFSKSSIEVKYSGTLLLSSNNS 538
          | |||:|||  :||:|:::| ||  || :||    ::||||| ||:|||||:||||:|
Db    481 KSLIKIRKEYGLACNRKLSVENGNYWLTVKGNGCLAVYVFSKSVIEMKYSGTLVLSSNSS 540

Qy    539 FPQHIEEGKYEFDKGFALYKL 559
          ||  | | ||| |||||||||
Db    541 FPSQITESKYELDKGFALYKL 561


RESULT  4
Q53641
ID   Q53641      PRELIMINARY;      PRT;   556 AA.
AC   Q53641; O08279; O08064;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   ALPHA-AMYLASE (FRAGMENT).
GN   TREZ.
OS   Sulfolobus acidocaldarius.
OC   Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.
OX   NCBI_TaxID=2285;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kato M., Kettoku M., Miura Y., Komeda T., Konishi Y., Shindo K.,
RA   Kobayashi K., Iwamatsu A.;
RL   Submitted (SEP-1995) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC33909;
RX   MEDLINE=97135071; PubMed=8980629;
RA   Maruta K., Mitsuzumi H., Nakada T., Kubota M., Chaen H., Fukuda S.,
RA   Sugimoto T., Kurimoto M.;
RT   "Cloning and sequencing of a cluster of genes encoding novel enzymes
RT   of trehalose biosynthesis from thermophilic archaebacterium Sulfolobus
RT   acidocaldarius.";
RL   Biochim. Biophys. Acta 1291:177-181(1996).
DR   EMBL; D64131; BAA11011.1; -.
DR   EMBL; D83245; BAA11863.1; -.
DR   InterPro; IPR000461; Alpha_amylase.
DR   InterPro; IPR004193; isoamylase_N.
DR   InterPro; IPR003662; sub_transporter.
DR   Pfam; PF00128; alpha-amylase; 1.
DR   Pfam; PF02922; isoamylase_N; 1.
DR   PROSITE; PS00217; SUGAR_TRANSPORT_2; UNKNOWN_1.
KW   Hydrolase.
FT   NON_TER     556     556
SQ   SEQUENCE   556 AA;  64373 MW;  4215B45E6C8ED4E7 CRC64;


  Query Match             58.7%;  Score 1748.5;  DB 1;  Length 556;
  Best Local Similarity   58.9%;  Pred. No. 8.9e-103;
  Matches  329;  Conservative   87;  Mismatches  136;  Indels    7;  Gaps    5;

Qy      2 TFAYKIDGNEVIFTLWAPYQKSVKLKVLEKGLYEMERDEKGYFTITLNNVKVRDRYKYVL 61
            :|   |: |: || |||||  ||||| | | |  ||::::|:| : :::::    | |::
Db      3 SFGGNIEKNKGIFKLWAPYVNSVKLK-LSKKLIPMEKNDEGFFEVEIDDIEENLTYSYII 61

Qy     62 DDASEIPDPASRYQPEGVHGPSQIIQESKEFNNETFLKKEDLIIYEIHVGTFTPEGTFEG 121
          :|  |||||||||| ||| ||:|:   :  :   :| |||||||:|||||: || |:|
Db     62 EDKREIPDPASRYQPLGVHDKSQLIRTDYQILDLGKVKIEDLIIYELHVGTFSQEGNFKG 121

Qy    122 VIRKLDYLKDLGITAIEIMPIAQFPGKRDWGYDGVYLYAVQNSYGGPEGFRKLVDEAHKK 181
          || |||||||||| ||:||:||||| |||||||:||||||:|||    |||:||||:
Db    122 VIEKLDYLKDLGITGIELMPVAQFPGNRDWGYDGVFLYAVQNTYGGPWELAKLVNEAHKR 181

Qy    182 GLGVILDVVYNHVGPEGNYMVKLGPYFSQKYKTPWGLTFNFDDAESDEVRKFILENVEYW 241
          |: |||||||||:||||||:: |||||| :|||||||||||||   |:|||||||||||
Db    182 GIAVILDVVYNHIGPEGNYLLGLGPYFSDRYKTPWGLTFNFDDRGCDQVRKFILENVEYW 241

Qy    242 IKEYNVDGFRLDAVHAIIDTSPKHILEEIADVVHKYNRIVIAESDLNDPRVVNPKEKCGY 301
          | : :|| |||||||| | ||||||:|||:  |:  : ||||||||||::|  |: |||
Db    242 FKTFKIDGLRLDAVHAIFDNSPKHILQEIAEKAHQLGKFVIAESDLNDPKIV--KDDCGY 299

Qy    302 NIDAQWVDDFHHSIHAYLTGERQGYYTDFGNLDDIVKSYKDVFVYDGKYSNFRRKTHGEP 361
          |||||||||||::||::| |:  || ||| ::|| |::||||||||||| :| :||| |
Db    300 KIDAQWVDDFHHAVHAFITKEKDYYYQDFGRIEDIEKTFKDVFVYDGKYSRYRGRTHGAP 359

Qy    362 VGELDGCNFVVYIQNHDQVGNRGKGERIIKLVDRESYKIAAALYLLSPYIPMIFMGEEYG 421
          ||:|    |||:|||||||||| |||:  | |: :| :|| ||:||||||:|||||||
Db    360 VGDLPPRKFVVFIQNHDQVGNRGNGERLSILTDKTTYLMAATLYILSPYIPLIFMGEEYY 419

Qy    422 EENPFYFFSDFSDSKLIQGVREGRKKENGQDTDPQDESTFNASKLSWKIDEEIFSFYKIL 481
          | |||:|||||||  ||:|||||| ||| |  ||| |  |  ||||||||||:  :|| |
Db    420 ETNPFFFFSDFSDPVLIKGVREGRLKENNQMIDPQSEEAFLKSKLSWKIDEEVLDYYKQL 479

Qy    482 IKMRKELSIACDRRVNVVNGENWLIIKGREYFSLYVFSKSSIEVKYSGTLLLSSNNSFPQ 541
          | :||  :  | |   |    | : :  :  :  |    |  | :| ||:    ||:
Db    480 INIRKRYN-NCKRVKEVRREGNCITLIMEKIGIIASFDDIVINSKITGNLLIGI--GFPK 536

Qy    542 HIEEGK-YEFDKGFALYKL 559
            ::: :  : ::|  :|:|
Db    537 KLKKDELIKVNRGVGVYQL 555


RESULT  5
Q973H4
ID   Q973H4      PRELIMINARY;      PRT;   359 AA.
AC   Q973H4;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PUTATIVE ALPHA-AMYLASE.
GN   ST0927.
OS   Sulfolobus tokodaii.
OC   Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.
OX   NCBI_TaxID=111955;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JCM 10545 / 7;
RX   PubMed=11572479;
RA   Kawarabayasi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
RA   Sekine M., Baba S.-I., Ankai A., Kosugi H., Hosoyama A., Fukui S.,
RA   Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
RA   Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,
RA   Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,
RA   Oshima T., Kikuchi H.;
```

```
RT   "Complete genome sequence of an aerobic thermoacidophilic
RT   Crenarchaeon, Sulfolobus tokodaii strain7.";
RL   DNA Res. 8:123-140(2001).
DR   EMBL; AP000984; BAB65939.1; -.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   359 AA;  41600 MW;  4D9F429099E81665 CRC64;
```

```
  Query Match           42.6%;  Score 1269;  DB 17;  Length 359;
  Best Local Similarity   64.9%;  Pred. No. 1.1e-72;
  Matches  233;  Conservative   51;  Mismatches   71;  Indels    4;  Gaps    3;

Qy     201 MVKLGPYFSQKYKTPWGLTFNFDDAESDEVRKFILENVEYWIKEYNVDGFRLDAVHAIID 260
           |: |||||||||||||||||||||| |||||||||:|| ||   ::||| |||||||| |
Db       1 MMFLGPYFSQKYKTPWGLTFNFDDAYSDEVRKFILDNVRYWFNVFHVDGLRLDAVHAIYD 60

Qy     261 TSPKHILEEIADVVHKYNRIVIAESDLNDPRVVNPKEKCGYNIDAQWVDDFHHSIHAYLT 320
           :||||||::||:: |   : |||||||||:::::  :|||| |||||||||||:|||||
Db      61 SSPKHILQDIAELSHSLGKFVIAESDLNDPKIID--DKCGYKIDAQWVDDFHHSVHAYLT 118

Qy     321 GERQGYYTDFGNLDDIVKSYKDVFVYDGKYSNFRRKTHGEPVGELDGCNFVVYIQNHDQV 380
           |||  ||:||| ||||||::|||||||||||||:|||| ||| |: | ||||||||||
Db     119 GERNSYYSDFGTLDDIVKAFKDVFVYDGKYSNFRKKTHGAPVGNLNACKFVVYIQNHDQV 178

Qy     381 GNRGKGERIIKLVDRESYKIAAALYLLSPYIPMIFMGEEYGEENPFYFFSDFSDSKLIQG 440
           |||| |:|:  |||::|| ||:|||||||:|||||||| |:||| |||||||  ||:|
Db     179 GNRGNGDRLSTLVDKQSYMIASALYLLSPFIPMIFMGEEYYEKNPFLFFSDFSDPNLIKG 238

Qy     441 VREGRKKENGQDTDPQDESTFNASKLSWKIDEEIFSFYKILIKMRKELSIACDRRVNVVN 500
           ||||| ::|||  ||| :  |  ||||| ||:|:  :|: |||:|||    | ||| |
Db     239 VREGRLRDNGQSIDPQSDEAFLKSKLSWNIDKEMLEYYRTLIKVRKEFGNKCSRRVEVDK 298

Qy     501 GENWLIIKGREYFSLYVFSKSSIEVKYSGTLLLSSNNSFPQHIEEGK-YEFDKGFALYK 558
           :|||:|   : | |  || : : ::|:| || |::: ||: |   |  :  ||  :||
Db     299 TDNWLVIGLEKIFLLVAFSDTGLILRYNGELLFSTSD-FPKRIVSSKEIKVKKGVGVYK 356
```

```
RESULT   6
Q9ADI5
ID   Q9ADI5      PRELIMINARY;      PRT;   581 AA.
AC   Q9ADI5;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE   PUTATIVE ALPHA AMYLASE.
GN   SCBAC1A6.02C.
OS   Streptomyces coelicolor.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=1902;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3(2);
RA   Seeger K.J., Harris D.;
RL   Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3(2);
RA   Cerdeno A.M., Parkhill J., Barrell B.G., Rajandream M.A.;
RL   Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3(2);
RX   MEDLINE=97000351; PubMed=8843436;
RA   Redenbach M., Kieser H.M., Denapaite D., Eichner A., Cullum J.,
RA   Kinashi H., Hopwood D.A.;
RT   "A set of ordered cosmids and a detailed genetic and physical map for
RT   the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL   Mol. Microbiol. 21:77-96(1996).
DR   EMBL; AL589708; CAC33923.1; -.
DR   InterPro; IPR000461; Alpha_amylase.
DR   Pfam; PF00128; alpha-amylase; 1.
SQ   SEQUENCE   581 AA;  63952 MW;  8816DFF3BC1A3408 CRC64;
```

```
  Query Match           34.7%;  Score 1033.5;  DB 2;  Length 581;
  Best Local Similarity   40.3%;  Pred. No. 1.7e-57;
  Matches  213;  Conservative   92;  Mismatches  185;  Indels   39;  Gaps   13;

Qy      14 FTLWAPYQKSVKLKVLEKGLYEMERDEK--GYFTITLNNVKVRD--RYKYVLDDASEIPD 69
           | :|||    | |:  :     :||| :  |::          ||  || : :||   :||
Db       3 FEVWAPQAGRVTLR-CDGATRALERDPERPGWW---CGEAPARDGSRYGFAVDDGPVLPD 58

Qy      70 PASRYQPEGVHGPSQIIQESKEFNNETFLKKEDL---IIYEIHVGTFTPEGTFEGVIRKL 126
           | || ||:|  | | ::   : :   |       |   ::||:||||:||||| :    :|
Db      59 PRSRRQPDGPDGLSAVVDHGR-YAWRTPWAGRPLPGGVLYELHVGTYTPEGTLDAAADRL 117

Qy     127 DYLKDLGITAIEIMPIAQFPGKRDWGYDGVYLYAVQNSYGGPEGFRKLVDEAHKKGLGVI 186
           ::|  ||:| :|:||:  |||:  |||:|| |:||   |||||  :: || ||  ||||:
Db     118 EHLVRLGVTHVELMPLCPFPGRHGWGYEGVSLWAVHEPYGGPEALKRFVDRAHDLGLGVV 177

Qy     187 LDVVYNHVGPEGNYMVKLGPYFSQKYKTPWGLTFNFDDAESDEVRKFILENVEYWIKEYN 246
           ||||:||:|| ||::   ||||:  : ||||:  | |   |||| ::::   |:::|
Db     178 LDVVHNHLGPSGNHLPAFGPYFTDTHHTPWGVAVNLDAPGSDEVRAYLVDSALAWLRDYR 237

Qy     247 VDGFRLDAVHAIIDTSPKHILEEIADVVH------KYNRIVIAESDLNDPRVVNPKEKCG 300
           :|| |||||||: ||   | ||::: |       :    :|||||||||:: |: : |
Db     238 LDGLRLDAVHALADTRACHFLEQLSTAVDGLAADLERPLFLIAESDLNDPRIITPRAEGG 297

Qy     301 YNIDAQWVDDFHHSIHAYLTGERQGYYTDFGN--LDDIVKSYKDVFVYDGKYSNFRRKTH 358
           : ||| |||||::|| |||| |||| ||    |  : |:    : :|| ||:|| ::|
Db     298 LGVHAQWNDDFHHALHAALTGESQGYYADFARDPLAALAKTLIRGYFHDGTYSSFRGRSH 357

Qy     359 GEPV--GELDGCNFVVYIQNHDQVGNRGKGERIIKLVDRESYKIAAALYLLSPYIPMIFM 416
           |  :  | :       | | ||||||| :|:|:  ||       || | | :|: ||:||
Db     358 GRALDRGRVAAHRLTGYSQTHDQVGNRAQGDRLASLVSPGLAACAATLTLTAPFTPMLFM 417

Qy     417 GEEYGEENPFYFFSDFSDSKLIQGVREGRKKE------NGQDT-DPQDESTFNASKLSWK 469
           |||:    |: ||:| :| :| | ||:||::|        :|  |||| :|   | | |
Db     418 GEEWAAGTPWQFFTDHTDPQLAQAVRQGRRREFAAHGWREEDVPDPQDPATRERSCLDWS 477

Qy     470 IDE-----EIFSFYKILIKMRKEL-----SIACDRRVNVVNGENWLIIK 508
             |      :  :|: || :| |          | ||       ||  :
Db     478 EPEREPHARVLDWYRRLIALRAEQPDLTDPDLADTRVAYDGTGRWLAFR 526
```

```
RESULT   7
Q9RX51
ID   Q9RX51      PRELIMINARY;      PRT;   600 AA.
AC   Q9RX51;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   MALTOOLIGOSYLTREHALOSE TREHALOHYDROLASE, PUTATIVE.
GN   DR0464.
OS   Deinococcus radiodurans.
OC   Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.
OX   NCBI_TaxID=1299;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=R1;
RX   MEDLINE=20036896; PubMed=10567266;
RA   White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA   Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA   Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA   Vamathevan J.J., Lam P., McDonald L., Utterback T., Zalewski C.,
RA   Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA   Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA   Fraser C.M.;
RT   "Genome sequence of the radioresistant bacterium Deinococcus
RT   radiodurans R1.";
RL   Science 286:1571-1577(1999).
DR   EMBL; AE001905; AAF10042.1; -.
DR   TIGR; DR0464; -.
DR   InterPro; IPR000461; Alpha_amylase.
DR   InterPro; IPR004193; isoamylase_N.
```

```
DR   Pfam; PF00128; alpha-amylase; 1.
DR   Pfam; PF02922; isoamylase_N; 1.
KW   Hydrolase; Complete proteome.
SQ   SEQUENCE   600 AA;  66909 MW;  594091EC093F8A44 CRC64;


  Query Match             30.4%;  Score 905.5;  DB 16;  Length 600;
  Best Local Similarity   38.7%;  Pred. No. 2.3e-49;
  Matches  205;  Conservative   86;  Mismatches  170;  Indels   69;  Gaps   20;

Qy      14 FTLWAPYQKSVKLKV---------LEKGLYEMERDEKGYFTITLNNVKVRDRYKYVLDDA 64
           | ||    ::| ::|         |  |:||:|             |    || :|||
Db      38 FRLWTSTARTVAVRVNGTEHVMTSLGGGIYELELP----------VGPGARYLFVLDGV 86

Qy      65 SEIPDPASRYQPEGVHGPSQIIQESKEFNNETF---------LKKEDLIIYEIHVGTFTP 115
              ||| :|: |:|||| ::::    :|   ||         :|  | : ||:|||||||
Db      87 -PTPDPYARFLPDGVHGEAEVV----DFG--TFDWTDADWHGIKLADCVFYEVHVGTFTP 139

Qy     116 EGTFEGVIRKLDYLKDLGITAIEIMPIAQFPGKRDWGYDGVYLYAVQNSYGGPEGFRKLV 175
           |||:     || |||:||:|||::||:| | |:| |||||   ||    || ||    ||
Db     140 EGTYRAAAEKLPYLKELGVTAIQVMPLAAFDGQRGWGYDGAAFYAPYAPYGRPEDLMALV 199

Qy     176 DEAHKKGLGVILDVVYNHVGPEGNYMVKLGP-YFSQKYKTPWGLTFNFDDAESDEVRKFI 234
           | ||: |||| |||||||| || |||:    | ||: :: : ||:    | ||   :|:::
Db     200 DAAHRLGLGVFLDVVYNHFGPSGNYLSSYAPSYFTDRFSSAWGM--GLDYAE-PHMRRYV 256

Qy     235 LENVEYWIKEYNVDGFRLDAVHAIIDTSPKHILEEIADVVHKY--NRIVIAESDLNDPRV 292
             |   |:::|: || ||||   : | |  ||| |:|  :|:     :::||   | | :
Db     257 TGNARMWLRDYHFDGLRLDATPYMTDDSETHILTELAQEIHELGGTHLLLAEDHRNLPDL 316

Qy     293 VNPKEKCGYNIDAQWVDDFHHSIHAYLTGERQGYYTDF-GNLDDIVKSYKDVFVYDGKYS 351
           |        ::|  | |||||     ||||::|||  : |  : :  :  :  : |:|::
Db     317 VTVN-----HLDGIWTDDFHHETRVTLTGEQEGYYAGYRGGAEALAYTIRRGWRYEGQFW 371

Qy     352 NFRRKTH--GEPVGELDGCNFVVYIQNHDQVGNRGKGERIIKL--VDRESYKIAAALYLL 407
             : : |  | |   |:  |||  ||||||:|||  |||: :   |    |: |||| |
Db     372 AVKGEEHERGHPSDALEAPNFVYCIQNHDQIGNRPLGERLHQSDGVTLHEYRGAAALLL- 430

Qy     408 SPYIPMIFMGEEYGEENPFYFFSDFSDSKLIQGVREGRKKE-------NGQDT-DPQDES 459
           |  |::| |:|:    || |||| :  :| | | ||||||       :|:|  ||| |
Db     431 -PMTPLLFQGQEWAASTPFQFFSDHA-GELGQAVSEGRKKEFGGFSGFSGEDVPDPQAEQ 488

Qy     460 TFNASKLSWKIDE-----EIFSFYKILIKMRKELSIACDR-RVNVVNGEN 503
           ||  |||:|   |          |: |:::|:|  :  :| | |:  | :
Db     489 TFLNSKLNWAEREGGEHARTLRLYRDLLRLRREDPVLHNRQRENLTTGHD 538


RESULT   8
Q9I1V1
ID   Q9I1V1      PRELIMINARY;      PRT;   583 AA.
AC   Q9I1V1;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE   PROBABLE GLYCOSYL HYDROLASE.
GN   PA2164.
OS   Pseudomonas aeruginosa.
OC   Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC   Pseudomonas.
OX   NCBI_TaxID=287;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 15692 / PAO1;
RX   MEDLINE=20437337; PubMed=10984043;
RA   Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrener P.,
RA   Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA   Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA   Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA   Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA   Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT   "Complete genome sequence of Pseudomonas aeruginosa PAO1, an
RT   opportunistic pathogen.";
RL   Nature 406:959-964(2000).

DR   EMBL; AE004643; AAG05552.1; -.
DR   InterPro; IPR002086; Aldehyde_dehydr.
DR   InterPro; IPR000461; Alpha_amylase.
DR   Pfam; PF00128; alpha-amylase; 1.
DR   PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
KW   Hydrolase; Complete proteome.
SQ   SEQUENCE   583 AA;  65657 MW;  651CD7A918833E00 CRC64;


  Query Match             29.6%;  Score 881.5;  DB 16;  Length 583;
  Best Local Similarity   36.3%;  Pred. No. 7.3e-48;
  Matches  203;  Conservative   95;  Mismatches  174;  Indels   87;  Gaps   19;

Qy       3 FAYKIDGN-EVIFTLWAPYQKSVKLKVLEKGLYEMERDEKGYFTITLNNVKVRDRYKYVL 61
           |  :  ||    | ||||  : |:::  :   : :|  : |:|  ||       ||:| :
Db       7 FGAQFQGNGRTCFGLWAPDAREVRVETADGRDWPLEGSDSGWFEATL-PCPPGTRYRYRI 65

Qy      62 DDASEIPDPASRYQPEGVHGPSQIIQES------KEFNNETFLKKEDLIIYEIHVGTFTP 115
           |    :|||||::||:|||| ||::          |:    :    : :|||:||| |
Db      66 DGRPGVPDPASQFQPDGVHGHSQVLDHGTYAWRVDEWRGRPW---HEAVIYELHVGLF-- 120

Qy     116 EGTFEGVIRKLDYLKDLGITAIEIMPIAQFPGKRDWGYDGVYLYAVQNSYGGPEGFRKLV 175
             |::  | | |  | :||:||:|:||: :|||:|:|||||  :|  ::|| ||  : |:
Db     121 -GSYAEVERFLPRLVELGVTAVELMPLGEFPGRRNWGYDGVLPFAPASAYGTPEQLKHLI 179

Qy     176 DEAHKKGLGVILDVVYNHVGPEGNYMVK-LGPYFSQKYKTPWGLTFNFDDAESDEVRKFI 234
           | ||  || | :||:||| ||:|||: :    :|    :||||   :|      |||:|
Db     180 DSAHGMGLMVFVDVIYNHFGPDGNYLAQYAAAFFRDDRQTPWGQAIDF---RRGEVREFF 236

Qy     235 LENVEYWIKEYNVDGFRLDAVHAIIDTSPKHILEEIA----------DVVHKYNRIVIAE 284
             ||   |: :| ||| | ||||||  |::    | |:|          ||     :: |
Db     237 YENALMWLLDYRVDGLRFDAVHAIPDSA---FLVEMARRLRGAAGPERHVH-----LVLE 288

Qy     285 SDLNDPRVVNPKEKCGYNIDAQWVDDFHHSIHAYLTGERQGYYTDFGN-LDDIVKSYKDV 343
           :| |   ::    : ||  |||| || ||::|  |||  ||| |:   |  : :   :
Db     289 NDDNRASLL----RQGY--DAQWNDDGHHALHVLLTGENDGYYQDYPEPLRCLARCLAEG 342

Qy     344 FVYDGKYSNFRRKTHGEPVGELDGCNFVVYIQNHDQVGNRGKGERIIKLVDRESYKIAAA 403
           |||  |: :|   :  |||  :|    ||:::|||||||||  |||:  | : :: ::| |
Db     343 FVYQGE-ANRHGRPRGEPSADLAPDAFVLFLQNHDQVGNRAFGERLSVLAEPQALRLAIA 401

Qy     404 LYLLSPYIPMIFMGEEYGEENPFYFFSDFSDSKLIQGVREGRKKENGQD---------- 452
           | ||:| ||::|||||     || :|:|    :|   |||||:|| |:
Db     402 LQLLAPMIPLLFMGEECAAREPFLYFTD-HQGELADAVREGRRKEFGEFGRFGEGATLAS 460

Qy     453 -TDPQDESTFNASK-------LSWKIDEEIFSFYKILIKMRKELSI---------- 490
            ||    ||  |:        :|:       ||: |:::| |  |
Db     461 LPDPNAVETFERSRPGLADCDPAWR------GFYRQLLEIRHEHLIPRLRGARSLGVTTI 514

Qy     491 ---ACDRRVNVVNGENWLI 506
             |   |  : :| :| |
Db     515 AGAALSARWRLGDGSDWRI 533


RESULT   9
Q92U63
ID   Q92U63      PRELIMINARY;      PRT;   601 AA.
AC   Q92U63;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PUTATIVE 1,4-ALPHA-GLUCAN BRANCHING ENZYME PROTEIN (EC
DE   2.4.1.18).
OS   Rhizobium meliloti (Sinorhizobium meliloti).
OG   Plasmid pSymB (megaplasmid 2).
OC   Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC   Rhizobiaceae; Sinorhizobium.
OX   NCBI_TaxID=382;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=1021;
RX   MEDLINE=21396508; PubMed=11481431;
RA   Finan T.M., Weidner S., Wong K., Buhrmester J., Chain P.,
```

```
RA   Vorhoelter F.J., Hernandez-Lucas I., Becker A., Cowie A., Gouzy J.,
RA   Golding B., Puehler A.;
RT   "The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-
RT   fixing endosymbiont Sinorhizobium meliloti.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:9889-9894(2001).
DR   EMBL; AL603646; CAC49676.1; -.
KW   Transferase; Glycosyltransferase; Plasmid; Hypothetical protein;
KW   Complete proteome.
SQ   SEQUENCE   601 AA;  66954 MW;  024DA322A7B72C2E CRC64;


  Query Match           29.4%;  Score 876.5;  DB 16;  Length 601;
  Best Local Similarity  35.3%;  Pred. No. 1.6e-47;
  Matches  207;  Conservative   99;  Mismatches  204;  Indels   77;  Gaps   19;

Qy       7 IDGNEVIFTLWAPYQKSVKLKVLEKGLYEMERDEKGYFTITLNNVKVRDRYKYVLDDASE 66
           || :   | |||| :: | | ||   :::|:  : |:| |||   |  :|| : | | :|
Db      19 IDSDTCRFRLWAPDEREVDL-VLGGAVHKMQSLDGGWFEITL-AAKAGERYCFRLADGTE 76

Qy      67 IPDPASRYQPEGVHGPSQIIQESK-EFNNETFLKK--EDLIIYEIHVGTFTPEGTFEGVI 123
           : |||| |     | | ::  :  |:   ::  :  |: :| |:||| |||||||   |
Db      77 VADPASSAQEREASGTSIVVDHAAYEWQTSSWRGRPWEEAVISELHVGCFTPEGTFRAAI 136

Qy     124 RKLDYLKDLGITAIEIMPIAQFPGKRDWGYDGVYLYAVQNSYGGPEGFRKLVDEAHKKGL 183
           :| :|   |||||||||:||||| | ||||||  ||  |:|| |:  : ||| ||  ||
Db     137 ERLPHLAGAGITAIEIMPVAQFPGVRGWGYDGVLHYAPHNAYGKPDDLKALVDAAHSLGL 196

Qy     184 GVILDVVYNHVGPEGNYMVKLGP-YFSQKYKTPWGLTFNFDDAESDEVRKFILENVEYWI 242
           |:||||||| ||||||: :    :|::   |||| :  |   | : ||:: :|| ||:
Db     197 TVLLDVVYNHFGPEGNYLSRYASRFFNKDRPTPWGASIAF---EEEAVRRYFIENALYWL 253

Qy     243 KEYNVDGFRLDAVHAIIDTSPKHILEEIADVVHK---YNRIVIAESDLNDPRVVNPKEKC 299
             :  || ||||   | ||:  | |  :   | :     :| :   | :  | :  ::
Db     254 GHFRFDGLRLDATEQIRDTTKPHFLVALEHEVREAFAERQIHLVLEDAHRRRSLLQRDAS 313

Qy     300 GYNI--DAQWVDDFHHSIHAYLTGERQGYYTDFGN--LDDIVKSYKDVFVYDGKYSNFR- 354
           |  :  || | || |:::|   ||| :|:|  | :     |  :  : |    |  ||
Db     314 GARMLFDAAWNDDLHNALHVVATGETKGHYRLFADEPWGKIRSALAEGFAVPAKEDNFSP 373

Qy     355 --RKTHGEPVGELDGCNFVVYIQNHDQVGNRGKGERIIKLVDRESYKIAAALYLLSPYIP 412
             :   | | ::      ::|||||:||| |||:  |:  :| :: ||:::|:| ||
Db     374 EGSRARVPPQGRVN------FLQNHDQIGNRAFGERLASLLQEDSLRVLAAMHMLTPQIP 427

Qy     413 MIFMGEEYGEENPFYFFSDFSDSKLIQGVREGRKKE--------NGQD----TDPQDEST 460
           ::|||||||  |||||||:   ::   :| ||: |         |:      || |
Db     428 LLFMGEEYGETQPFYFFSDY-QGEIAAAIRLGRRDEAENFGGLPEGKTMDDLPDPLDPDV 486

Qy     461 FNASKLSWK-----IDEEIFSFYKILIKMRK---------------------------- 486
           |  ||| |        |   :: : |  :|:
Db     487 FAGSKLRWNRATSPAGERHLAYVRDLAVIRQRHIVPLIAGTAVPDGRAYETKDGIIAVDW 546

Qy     487 ELSIAC-DRRVNVVNGENWL-IIKGREYFSLYVFSKSSIEVKYSGTL 531
           :   || : |||:::  : :  |:|:  |:      ||:|  |   |
Db     547 QFGEACLELRVNLLHETHAVPAIRGQPIFT---SETSSVETVYGSEL 590


RESULT  10
Q973H5
ID   Q973H5      PRELIMINARY;      PRT;   217 AA.
AC   Q973H5;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PUTATIVE ALPHA-AMYLASE.
GN   ST0926.
OS   Sulfolobus tokodaii.
OC   Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.
OX   NCBI_TaxID=111955;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JCM 10545 / 7;
RX   PubMed=11572479;
RA   Kawarabayasi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
RA   Sekine M., Baba S.-I., Ankai A., Kosugi H., Hosoyama A., Fukui S.,
RA   Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
RA   Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,
RA   Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,
RA   Oshima T., Kikuchi H.;
RT   "Complete genome sequence of an aerobic thermoacidophilic
RT   Crenarchaeon, Sulfolobus tokodaii strain7.";
RL   DNA Res. 8:123-140(2001).
DR   EMBL; AP000984; BAB65938.1; -.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   217 AA;  25181 MW;  4A1235EC5B491CEF CRC64;


  Query Match           23.7%;  Score 706;  DB 17;  Length 217;
  Best Local Similarity  66.0%;  Pred. No. 2.4e-37;
  Matches  132;  Conservative   28;  Mismatches   36;  Indels    4;  Gaps    4;

Qy      10 NEVIFTLWAPYQKSVKLKVLEKGLYEMERDEKGYFTITLNNVKVRDRYKYVLDDASEIPD 69
           ||| | |||||||   |:: |:  |:||:|||||| ||: | |  |:| ::: |  ||||
Db      10 NEVEFILWAPYQKKATLRLNEE-KYDMEKDEKGYFRITV-NAKEGDKYSFII-DKGEIPD 66

Qy      70 PASRYQPEGVHGPSQIIQESKEF-NNETFLKKEDLIIYEIHVGTFTPEGTFEGVIRKLDY 128
           |||||||:|||| ||:|    |: : |  :| ||||||||:|:|||| ||||   : ||||
Db      67 PASRYQPDGVHGRSQLISLKYEWKSREVKIKPEDLIIYELHIGTFTQEGTFYSAVEKLDY 126

Qy     129 LKDLGITAIEIMPIAQFPGKRDWGYDGVYLYAVQNSYGGPEGFRKLVDEAHKKGLGVILD 188
           | :||:|||||||: |||||:|||||||||||||||:|||| | | :||||::|| ||||
Db     127 LVNLGVTAIEIMPVHQFPGKKDWGYDGVYLYAVQNTYGGPYAFMKFIDEAHQRGLAVILD 186

Qy     189 VVYNHVGPEGNYMVKLGPYF 208
           |||||||||||||: ||||:
Db     187 VVYNHVGPEGNYMMFLGPYY 206


RESULT  11
Q44528
ID   Q44528      PRELIMINARY;      PRT;   552 AA.
AC   Q44528;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PUTATIVE  ALPHA-GLUCANOTRANSFERASE /HYDROLASE.
GN   LTI2.
OS   Anabaena variabilis.
OC   Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Anabaena.
OX   NCBI_TaxID=1172;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=M3;
RX   MEDLINE=92119230; PubMed=1731972;
RA   Sato N.;
RT   "Cloning of a low-temperature-induced gene lti2 from the
RT   cyanobacterium Anabaena variabilis M3 that is homologous to alpha-
RT   amylases.";
RL   Plant Mol. Biol. 18:165-170(1992).
DR   EMBL; X59279; CAA41970.1; -.
DR   InterPro; IPR000461; Alpha_amylase.
DR   InterPro; IPR004193; isoamylase_N.
DR   Pfam; PF00128; alpha-amylase; 1.
DR   Pfam; PF02922; isoamylase_N; 1.
KW   Transferase; Hydrolase.
SQ   SEQUENCE   552 AA;  63833 MW;  069886FAFD8FC411 CRC64;


  Query Match           17.6%;  Score 524.5;  DB 2;  Length 552;
  Best Local Similarity  28.0%;  Pred. No. 2.7e-25;
  Matches  148;  Conservative   90;  Mismatches  162;  Indels  129;  Gaps   22;

Qy      14 FTLWAPYQKSVKLKVLEKGLYE--MERDEKGYFTITLNNVKVRD---RYKYVLDDAS--- 65
           |||:||| |   |         |  |:: : |||  |   |:: |   :||: :   |
Db       7 FTLFAPYNKGAALIASFSDWQEIPMKKGDDGYFRTT---VELEDGTYQYKFRVQTRSWFF 63

Qy      66 ------EIPDPASRYQPEGVHGPSQIIQESKEFNN--ETFLKKED---------LIIYEI 108
```

```
                 :: || :    |   |    |   |:     :|:: : |         |:|||:
Db     64 EEDQWVDVTDPYATDIDES-SGKDNSIARIKDGEKIVDTYVWQHDDKPLPADHELVIYEL 122

Qy    109 HVGTFT-------PEGTFEGVIRKLDYLKDLGITAIEIMPIAQFPGKRDWGYDGVYLYAV 161
          ||| |:          | :: || ||||| :||| |||::|: ::||   |||:  | :|
Db    123 HVGDFSGGEDDPYARGKYKHVIEKLDYLCELGINAIELLPVKEYPGDYSWGYNPRYFFAT 182

Qy    162 QNSYGGPEGFRKLVDEAHKKGLGVILDVVYNHVGPEGNY-MVKLGPYFSQKYKTP---WG 217
          ::|||     :|||| |::|: :|:| :|||         :    ::  : : |   ||
Db    183 ESSYGSTADLKKLVDECHQRGIRIIMDGIYNHSEASSPLTQIDHDYWYHHEPRDPDNNWG 242

Qy    218 LTFNF----DDAESDEVRKFILENVEYWIKEYNVDGFRLDAVHAIIDTSPKHIL----EE 269
            ||:    :: |:   |||| : | ||: ||::|| | ||   | :   | :    ::
Db    243 PEFNYEHYDENLETYPARKFIGDTVRYWVGEYHLDGIRYDAARQIANYDFMHWIAQEAKK 302

Qy    270 IADVVHKYNRIVIAESDLNDPRVVN---PKEKCGYNIDAQWVDDFHHSIHAYLTGERQGY 326
            |     ||   :||       : |   | : |       | | |:|:| |:: |:
Db    303 TAGAKPFYN---VAEHIPETTSITNLDGPMDGC-------WHDSFYHTIKAHICGDT--- 349

Qy    327 YTDFGNLDDIVKSYKDVFVYDGKYSNFRRKTHGEPVGELDGCNFVVYIQNHD------QV 380
            |  || |::          | |   |            |   | | |: |||      ::
Db    350 -FDLENLKDVI---------DPKRQGF-----------LGATNVVNYLTNHDHDHIMVEL 388

Qy    381 GNRGKGERIIKLVDRESY---KIAAALYLLSPYIPMIFMGEEYGEENPFYFFSDFSDSKL 437
          |||       ::   |::   |:  |: : :  :|:|:||||:||  |
Db    389 GNR-------EIFHDEAFRRAKLGTAILMTAVGVPLIWMGEEFGEYKP------------ 429

Qy    438 IQGVREGRKKENGQDTDPQDESTFNASKLSWKIDEEIFSFYKILIKMRK 486
                        ||:|  : : |   ::  :| ::| || :||
Db    430 ----------------KQQDQSKIDWTLLGNDLNRSLFDYHKGLIGLRK 462


RESULT 12
Q59242
ID   Q59242      PRELIMINARY;      PRT;   652 AA.
AC   Q59242; O08511; O08485;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   1,4-ALPHA-GLUCAN BRANCHING ENZYME (EC 2.4.1.18)
DE   (GLYCOGEN BRANCHING ENZYME) (1,4-ALPHA-GLUCAN BRANCHING ENZYME)
DE   (AMYLO-(1,4 TO 1,6)TRANSGLUCOSIDASE)
DE   (AMYLO-(1,4-1,6)-TRANSGLYCOSYLASE).
GN   GLGB.
OS   Bacillus stearothermophilus.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Bacillus/Staphylococcus group; Geobacillus.
OX   NCBI_TaxID=1422;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TRBE14;
RX   MEDLINE=95031021; PubMed=7944355;
RA   Takata H., Takaha T., Kuriki T., Okada S., Takagi M., Imanaka T.;
RT   "Properties and active center of the thermostable branching enzyme
RT   from Bacillus stearothermophilus.";
RL   Appl. Environ. Microbiol. 60:3096-3104(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TRBE14;
RA   Takata H.;
RL   Submitted (AUG-1996) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TRBE14;
RA   Takata H., Takaha T., Okada S., Takagi M., Imanaka T.;
RL   Submitted (APR-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- CATALYTIC ACTIVITY: FORMATION OF 1,6-GLUCOSIDIC LINKAGES OF
CC       GLYCOGEN.
CC   -!- PATHWAY: THIRD STEP IN GLYCOGEN BIOSYNTHESIS.
CC   -!- MISCELLANEOUS: OPTIMAL ACTIVITY AT APPROXIMATELY 39 DEGREES
CC       CELSIUS.
CC   -!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC       KNOWN AS THE ALPHA-AMYLASE FAMILY.
DR   EMBL; D87026; BAA19588.1; -.
DR   InterPro; IPR000461; Alpha_amylase.
DR   InterPro; IPR004193; isoamylase_N.
DR   Pfam; PF00128; alpha-amylase; 1.
DR   Pfam; PF02922; isoamylase_N; 1.
KW   Glycogen biosynthesis; Transferase; Glycosyltransferase.
FT   ACT_SITE    308    308       BY SIMILARITY.
FT   ACT_SITE    351    351       BY SIMILARITY.
FT   ACT_SITE    419    419       BY SIMILARITY.
SQ   SEQUENCE   652 AA;  76793 MW;  4591EB414A0E3FEF CRC64;


  Query Match             13.4%;  Score 399;  DB 2;  Length 652;
  Best Local Similarity   25.2%;  Pred. No. 3e-17;
  Matches  149;  Conservative   96;  Mismatches  189;  Indels  158;  Gaps   31;

Qy      4 AYKIDGNEVI---------FTLWAPYQKSVKL----KVLEKGLYEMER-DEKGYFTITLN 49
            :|:: |  ||         | :|||: : |:|       :   : :|: :::| :|| :
Db     22 SYELFGAHVINEGGKVGTRFCVWAPHAREVRLVGSFNDWDGTDFRLEKVNDEGVWTIVVP 81

Qy     50 NVKVRDRYKY--VLDDASEI--PDPASRYQPEGVHGPSQIIQESK--EFNNETFLKK--- 100
              |||  |  |   :  || : |     |  | |  : |  ::|:::: :|
Db     82 ENLEGHLYKYEIVTPDGQVLFKADPYAFYSELRPHTAS-IAYDLKGYQWNDQSWKRKKRR 140

Qy    101 -----EDLIIYEIHVGTF-TPEGTF----EGVIRKLDYLKDLGITAIEIMPIAQFPGKRD 150
               : ::|||:| |::    :| |     |     : |: | | | ||::|: : |  |
Db    141 KRIYDQPMVIYELHFGSWKKKDGRFYTYREMADELISYVLDHGFTHIELLPLVEHPLDRS 200

Qy    151 WGYDGVYLYAVQNSYGGPEGFRKLVDEAHKKGLGVILDVVYNHV--GPEGNYMVKLGP-- 206
          ||| |   ||| : || | |   ||  |: |:|||:| |  |       | ||    |
Db    201 WGYQGTGYYAVTSRYGTPHDFMYFVDRCHQAGIGVIMDWVPGHFCKDAHGLYMFDGAPTY 260

Qy    207 -YFSQKYKT--PWGLTFNFDDAESDEVRKFILENVEYWIKEYNVDGFRLDAVHAII---- 259
           | ::| :    || | || |    ||| |:: |  :|:: |::|||:|||  ::
Db    261 EYANEKDRENYVWG-TANF-DLGKPEVRSFLISNALFWLEYYHIDGFRVDAVANMLYWPN 318

Qy    260 ------DTSPKHILEEIADVVHKY--NRIVIAESDLNDPRVVNPKEKCGYNIDAQWVDDF 311
                :      | :: : |  |  | ::|||   : |:|  |  : |   : :|
Db    319 NDRLYENPYAVEFLRKLNEAVFAYDPNALMIAEDSTDWPKVTAPTYEGGLGFNYKW---- 374

Qy    312 HHSIHAYLTGERQGYYTDFGNLDDIVKSYKDVFVYDGKYSNFRRKTHGEPVGELDGCNFV 371
                        : | ::|::| | :   |:      ||  | :    |
Db    375 -----------------NMGWMNDMLK-YMETPPYE------RRHVHNQVTFSL----LY 406

Qy    372 VYIQN------HDQVGNRGKGERIIKLVDRESYKIAAALYLLSPYI------PMIFMGEE 419
            | :|      ||:| : ||   : |:      |  | | ||  |:       ::||| |
Db    407 AYSENFILPFSHDEVVH-GKKSLLNKMPGSYEEKF-AQLRLLYGYMMAHPGKKLLFMGNE 464

Qy    420 YGEENPFYFFSDFSDSKLIQGVREGRKKENGQDTDPQDESTFNASKLSWKIDEEI----- 474
                    |: | :                                  || ::|:
Db    465 ---------FAQFDE--------------------------------WKFEDELDWVLF 482

Qy    475 -FSFYKILIKMRKELSIACDRRVNVV-------NGENWLIIKGREYFSLYVF 518
           |  :: :    ||| ||| :|           |  |: :   |  |:: |
Db    483 DFELHRKMNDYMKEL-IACYKRYKPFYELDHDPQGFEWIDVHNAEQ-SIFSF 532


RESULT 13
Q93Q35
ID   Q93Q35      PRELIMINARY;      PRT;   440 AA.
AC   Q93Q35;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   BRANCHING ENZYME GLGB (FRAGMENT).
OS   Myxococcus xanthus.
OC   Bacteria; Proteobacteria; delta subdivision; Myxobacteria;
OC   Myxococcales; Cystobacterineae; Myxococcaceae; Myxococcus.
OX   NCBI_TaxID=34;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Ueki T., Inouye S.;
RT   "Identification of a new His-Asp phosphorelay signal transduction
```

```
RT   system which regulates expression of a heat shock gene, lonD, of
RT   Myxococcus xanthus.";
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF285783; AAK83002.1; -.
FT   NON_TER     440    440
SQ   SEQUENCE   440 AA;  48964 MW;  8D492EA5A4A92017 CRC64;
```

```
  Query Match           13.3%;  Score 397;  DB 2;  Length 440;
  Best Local Similarity   28.0%;  Pred. No. 2.3e-17;
  Matches  131;  Conservative   75;  Mismatches  164;  Indels   98;  Gaps   23;

Qy      14 FTLWAPYQKSVKLKVLEKGLY-------EMERDEKGYFTITLNNVKVRDRYKYVL----- 61
           | :|||    :  :|   | :       |:  :  | |:  :      :||::
Db      12 FRVWAP----MASRVFVSGDFNGWGTWIELGNEFNGNFSGDVAGAVKGQKYKFITRNQWG 67

Qy      62 DDASEIPDPASRYQPEGVHGPSQIIQESKEF--NNETFLKK--EDLIIYEIHVGTF--TP 115
             ||    || | :|       | || :  |:  | : :      ::||||:|||||  :|
Db      68 SDAWR-ADPRSAWQENST--GSSIIYDHGEYWWNAQQYSSPGFNEMIIYEMHVGTFHDSP 124

Qy     116 ---EGTFEGVIRKLDYLKDLGITAIEIMPIAQFPGKRDWGYDGVYLYAVQNSYGGPEGFR 172
               | :   | :||:::|||   |::||  :| |   |||:  : :| :::|| |   :
Db     125 GFGPGNWNSAIARLDHVRDLGANMIKVMPAYEFAGDFSWGYNAAFPFAPESAYGHPNDMK 184

Qy     173 KLVDEAHKKGLGVILDVVYNHVGP--------EGNYMVKLGPYFSQKYK--TPWGLTFNF 222
           : ||||| :|:||| |||:|| ||          |: :   | ||   ::  ||||
Db     185 RFVDEAHMRGIGVIFDVVHNHYGPSDLPMWCFSGDCLGSGGEYFYNDWRKSTPWGDT--R 242

Qy     223 DDAESDEVRKFILENVEYWIKEYNVDGFRLDAVHAI-----IDTSPKHILEEIADVVHKY 277
             |    ||| :| :::      :  || | ||   :      ||:    : :   |
Db     243 PDYGRPEVRAYIRDSMMNLTHSFRGDGLRWDATKYMRTQNGSDTT---AIPDAWRVFRSI 299

Qy     278 NR---------IVIAESDLNDPRVVN---PKEKCGYNIDAQWVDDFHHSIHAYLTGERQG 325
           ||         | |||       : |       |   |:||  || |:| | :
Db     300 NREINATQPWKISIAEDFGGGDFITNDATSDTSGGAGFDSQWGGDFVHAIRAAVIAS--- 356

Qy     326 YYTDFGNLDDIVKSYKDVFV--YDGKYSNFRRKTHGEPVGELDGCNFVVYIQNHDQVGNR 383
             | |     : | ::     | |:::                 |:| ::||:| |
Db     357 --NDSGR---DMNSVRNAITQRYSGRHT----------------ARVIYSESHDEVAN- 393

Qy     384 GKGERIIKLV---------DRESYKIAAALYLLSPYIPMIFMGEEYGE 422
           ||  |: : :          ::   :|| :   || ||||| |:|: |
Db     394 GKA-RVPEEIWPGNAGSWAAKKRSTLAAGVVFTSPGIPMIFQGQEFLE 440
```

```
RESULT  14
Q93UZ9
ID   Q93UZ9      PRELIMINARY;      PRT;  1142 AA.
AC   Q93UZ9;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   ALKALINE PULLULANASE.
GN   PUL1876.
OS   Bacillus sp. KSM-1876.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Bacillus/Staphylococcus group; Bacillus.
OX   NCBI_TaxID=72413;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=KSM-1876;
RA   Hatada Y., Ito S.;
RT   "alkaline pullulanase from the alkaliphilic bacterium.";
RL   Submitted (OCT-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB049812; BAB47586.1; -.
SQ   SEQUENCE   1142 AA;  128759 MW;  318FC01362961E84 CRC64;
```

```
  Query Match           13.3%;  Score 395.5;  DB 2;  Length 1142;
  Best Local Similarity   22.9%;  Pred. No. 1.1e-16;
  Matches  164;  Conservative  100;  Mismatches  190;  Indels  261;  Gaps   34;

Qy       8 DGNEVIFTLWAP---------YQK--------SVKLKVLEKGLYEMERDE--------KG 42
```

```
           ||   |  ||:|         | |         |::|: ::|::|::  :        :|
Db     376 DGTATI-KLWSPKAEHVQVILYDKDNPDDIVTEVEMKLGDRGVWEVQLTKENTGLDSLRG 434

Qy      43 YFTITLNNVKVRDRYKYVLD---DASEIPDPASR-----------YQPEGVHGPSQIIQ 87
           |:            | | :|   |     || ::           |    :  || |
Db     435 YY-----------YHYEIDHDGDKKIALDPYAKSLSAWNSDEQGPYAKAAIVDPSSIGP 482

Qy      88 ESKEFNNETFLKKEDLIIYEIHVGTFTPE----------GTFEGVIRKLDYLKDLGITA 136
           | :  : | | |||| ||||:||  || :           |||   : ||||::|||:|
Db     483 ELEFAHIEGFEKKEDTIIYEVHVRDFTSDPHIADELTAQFGTFASFVDKLDYIEDLGVTH 542

Qy     137 IEIMPIAQF----PGKRD-------------WGYDGVYLYAVQNSYG----GPE----G 170
           |:::|:  :      | |              ||||    :|:   |      ||
Db     543 IQLLPVMSYFFANEFKNDERMLDFSSTNTNYWGYDPQSYFALTGMYSEDPTDPELRIKE 602

Qy     171 FRKLVDEAHKKGLGVILDVVYNHVGPEGNYMVKLGPYFSQKYKTPWGLTFNFDDAES--- 227
           |:||:|| | :|:||:|||||||    | :   :  |            ::| ||:
Db     603 FKKLIDEIHSRGMGVVLDVVYNHTARVGIFEDLVPNY-----------YHFMDADGTPR 650

Qy     228 -----------DEVRKFILENVEYWIKEYNVDGFRLDAVHAIIDTSPKHILEEIADVVH 275
                       :  |: :::::  ||::|: :|||| | :          |  | |
Db     651 TSFGGGRLGTTHEMSRRILVDSITYWVEEFKIDGFRFDMM-------GDHDAESIQIAYD 703

Qy     276 KY-----NRIVIAE------SDLNDPRVVNPKEKCGYNIDAQWV----------DDFHHS 314
           |      | ::| |       | :|| |           |  |:          |:| :
Db     704 KAKEINPNIVMIGEGWITYAGDEDDPNV--------QAADQHWMQYTESVGVFSDEFRNE 755

Qy     315 IHA---------YLTGERQGYYTDFGNLD---DIVKSYKDVFVYDGKYSNFRRKTHGEPV 362
           : :          :|||  :       |:|   | :|:    |: |
Db     756 LKSGFGHEGEPRFLTGGAR-------NIDLIFDNIKAQPHNFIAD--------------- 793

Qy     363 GELDGCNFVVYIQNHDQVGNRGKGERIIKLVDR----------ESYKIAAALYLLSPYIP 412
             |  : | ||: ||   |    : |   : :           :  ::   : | |  |
Db     794 ---DPGDVVPYIEAHD---NLTLHDVIAMAIQKDPDDHQEEIHQRIRLGNLMTLTSQGIA 847

Qy     413 MIFMGEEYG--------------------EEN--PF---YFFSDFSDSKLIQGVREGRKK 447
           :  |:|||                    :||  ||   ||  |  ||  |    :  :
Db     848 FLHAGQEYGRTKQFRAETSEPPYKSTYMTDENGEPFRYPYFIHDSYDSTDIINRFDWERA 907

Qy     448 ENGQDTDPQDEST-------------FNASKLSWK--IDEEIFSFYKILIKMRKELSIAC 492
             |  |  |    |             :| :|  |  :||::    :: |   :|  :
Db     908 TNA-DAYPIQNLTREYTTGLIELRRSSDAFRLGTKDLVDEKV---TQLNIPEIEETDLVV 963

Qy     493 DRRVNVVNGENWLIIKGREYFSLYVFSKSSIEVKYSGTLLLSSNNSFPQHIEEGK 547
             |:    ||            : |||  :  | :   || |  : :  : : :||
Db     964 AYRIEATTGE-----------AFYVFVNADDEER---TLTLEEDLTVGEFVVDGK 1004
```

```
RESULT  15
Q59319
ID   Q59319      PRELIMINARY;      PRT;   825 AA.
AC   Q59319;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   ALPHA-DEXTRIN 6-GLUCANOHYDROLASE (EC 3.2.1.41).
GN   PULA.
OS   Caldocellum saccharolyticum (Caldicellulosiruptor saccharolyticus).
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Thermoanaerobacter group; Caldicellulosiruptor.
OX   NCBI_TaxID=44001;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Albertson G.D., McHale R., Gibbs M.D., Bergquist P.L.;
RT   "Cloning and expression of a type II pullulanase from an extremely
RT   thermophilic anaerobic bacterium, Caldicellulosiruptor
RT   saccharolyticus.";
RL   Eur. J. Biochem. 0:0-0(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Gibbs M.D.;
RL   Submitted (AUG-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; L39876; AAB06264.1; -.
```

```
DR   InterPro; IPR000461; Alpha_amylase.
DR   InterPro; IPR004193; isoamylase_N.
DR   Pfam; PF00128; alpha-amylase; 1.
DR   Pfam; PF02922; isoamylase_N; 1.
KW   Hydrolase; Glycosidase.
SQ   SEQUENCE   825 AA;  95732 MW;  92C2082E9B35F5AD CRC64;


  Query Match           12.8%;  Score 381.5;  DB 2;  Length 825;
  Best Local Similarity  25.7%;  Pred. No. 5.2e-16;
  Matches  149;  Conservative   86;  Mismatches  170;  Indels  175;  Gaps   31;

Qy      14 FTLWAPYQKSVKLKVLE-----KGLYEMERDEKGYFTITL-----NNVKVRDRYKYVLDD 63
           | ||||    ||::: :     |   || | | | : : |     |:  : : : |  :|
Db     220 FRLWAPTAYKVKIQIFDEHENFKFGKEMSRAENGTWDVYLPGDLKNHFYLYEIWHYNYED 279

Qy      64 -----ASEIPDPASRYQPEGVHGPSQIIQES----KEFNNETFL----KKEDLIIYEIHV 110
                 :||| |:       | | |   :      :  ::|:    |::| ||||:||
Db     280 DEGYIVYHVPDPYSKASSSN-SGKSYIFDPADGLIDGWQADSFVDNIEKQDDAIIYEMHV 338

Qy     111 GTFT------------------PEGTF-EGVIRKLDYLKDLGITAIEIMPIAQF----- 145
             ||                    || : ||:   | :||:||:| : ::||: |
Db     339 RDFTIDQSSGIGENLRGKFLGFCQEGYYKEGISTGLLHLKELGVTHVHLLPISDFGSVDD 398

Qy     146 --PGKR-DWGYDGV------YLYAVQNSYGGPEGFRKL---VDEAHKKGLGVILDVVYNH 193
             | || :|||| |      | |: ::  || |  ::|   : : |: |:||::|||:||
Db     399 KNPDKRYNWGYDPVLYQCPEYWYSTKS--GGIEALKELRTMIKKLHENGIGVVMDVVFNH 456

Qy     194 --------------VGPEGNYMVKLGPYFSQKYKTPWGLTFNFDDAESDEVRKFILENVE 239
                        : ||  |  ::  |      |  |   |    |   ||||||: :
Db     457 TYHTKGGKFSIFDKIVPE--YFYRVDDYGDYSNATGCG---NELATEKPMVRKFILDTII 511

Qy     240 YWIKEYNVDGFRLDAVHAIIDTSPKHILEEIADVVHKYN--------RIVIAESDLNDPR 291
           || :|:::|||| | :  :     :    ::|: | | |         |:  |
Db     512 YWTEEFHIDGFRFDLMGLLTVKQSR----QVANEVRKRNPKALVYGEGWVMGNSTCLVEE 567

Qy     292 VVNPKEKC--GYNIDAQWVDDFHHSIHAYLTGERQGYYTDF--GNLDDI------VKSYK 341
           :      |  ||:|       |:  |   : |:  || | :  ||| |:      :|:
Db     568 MATILSSCHQGYSIGL-----FNDRIRDAIRGDLDGYKTGYVHGNLSDVGRLKQGIKAAI 622

Qy     342 DVFVYDGKYSNFRRKTHGEPVGELDGCNFVVYIQNHDQVGNRGK------GERIIKLVDR 395
           | |                   | | |  | |:  || :    |      || |   :||
Db     623 DDF----------------AKEPDEC--VNYVSCHDNLTLFDKLQKTMVGEDIF-WIDR 662

Qy     396 ESYKIAAALYLLSPYIPMIFMGEEY-----GEENPFYFFSDFSDSKLIQGVREGRKKENG 450
            : ::| |: | |  :  :  | |:     |  | :                    |
Db     663 AT-RLANAIVLTSQGVAFLHGGVEFNRSKGGHPNTY----------------------NA 699

Qy     451 QDTDPQDESTFNASKLSWKIDE---EIFSFYKILIKMRKE 487
            |         | :|: | : |   : | ||  || :|::
Db     700 GD---------NINKIDWSLKEKFFDTFKFYCDLINLRRK 730


Search completed: July 15, 2002, 12:24:08
Job time: 390 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: July 15, 2002, 12:17:32 ; Search time 58.93 Seconds
(without alignments)
1047.973 Million cell updates/sec

Title: US-09-298-924-8
Perfect score: 2954
Sequence: 1 MFSFGGNIEKNKGIFKLWAP.....KLKKDELIKVNRGVGVYQLE 556

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802:*
1: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2954	100.0	556	17	AAR92755	Trehalose-releasin
2	2954	100.0	556	17	AAR90620	Sulfolobus acidoca
3	1748.5	59.2	559	17	AAR90619	Sulfolobus solfata
4	1042.5	35.3	595	22	AAG92072	C glutamicum prote
5	1042.5	35.3	610	22	AAB80206	Corynebacterium gl
6	969	32.8	589	16	AAR80289	Trehalose releasin
7	969	32.8	596	16	AAR77470	Trehalose releasin
8	959	32.5	598	16	AAR77471	Trehalose releasin
9	958	32.4	597	16	AAR80290	Trehalose releasin
10	936	31.7	575	21	AAY85157	Trehalose-releasin
11	936	31.7	575	21	AAY85165	Trehalose-releasin
12	360	12.2	652	16	AAR80037	Bacillus stearothe
13	360	12.2	652	17	AAR96109	Starch-branching-e
14	356	12.1	658	13	AAR23787	Heat-resistant pul
15	342.5	11.6	639	12	AAR11271	B.stearothermophil
16	336	11.4	731	22	AAG91100	C glutamicum prote
17	336	11.4	731	22	AAB79423	Corynebacterium gl
18	326.5	11.1	648	10	AAP94635	B. thuringiensis a
19	323	10.9	772	18	AAW34567	Thermotoga maritim
20	321	10.9	772	19	AAW49871	Thermotoga maritim
21	320	10.8	793	20	AAY00869	S. tuberosum isoam
22	313.5	10.6	1250	21	AAY91279	Group B Streptococ
23	312	10.6	670	22	AAU60913	Propionibacterium
24	310.5	10.5	606	18	AAW36602	S. tuberosum debra
25	306	10.4	718	22	AAU60903	Propionibacterium
26	306	10.4	766	20	AAY00870	S. tuberosum isoam
27	303.5	10.3	666	20	AAY34991	C. pneumoniae prot
28	301	10.2	562	10	AAP91904	Sequence of amy B
29	298.5	10.1	621	21	AAB19285	A polypeptide with
30	296.5	10.0	818	20	AAY17523	Zea mays SU1 starc
31	296	10.0	630	22	AAB69074	Aquifex aeolicus V
32	293.5	9.9	931	20	AAY27357	Group B Streptococ
33	292	9.9	829	22	AAE05691	Bacillus acidopull
34	292	9.9	862	21	AAY78513	Pullulanase (pulB)
35	285	9.6	764	21	AAY50819	Wheat isoamylase p
36	284	9.6	798	20	AAW73552	Pullulanase protei
37	284	9.6	1938	20	AAW73553	Full length Pullul
38	282	9.5	893	17	AAW09257	Bacillus alkaline
39	282	9.5	1938	17	AAW09255	Bacillus alkaline
40	280.5	9.5	931	20	AAY00872	Original S. tubero
41	278.5	9.4	618	20	AAY37184	Protein involved i
42	277	9.4	921	22	AAE05723	Bacillus acidopull
43	277	9.4	921	22	AAE05726	Bacillus acidopull
44	276	9.3	762	21	AAW90977	N. denitrificans a
45	276	9.3	921	22	AAE05689	Bacillus acidopull

ALIGNMENTS

RESULT 1
AAR92755
ID AAR92755 standard; Protein; 556 AA.
XX
AC AAR92755;
XX
DT 03-AUG-1996 (first entry)
XX
DE Trehalose-releasing thermostable enzyme.
XX
KW Thermostable enzyme; trehalose; sweetener.
XX
OS Sulfolobus acidocaldarius strain ATCC 33909.
XX
PN AU9527131-A.
XX
PD 01-FEB-1996.
XX
PF 21-JUL-1995; 95AU-0027131.
XX
PR 04-JUL-1995; 95JP-0189760.
PR 21-JUL-1994; 94JP-0190180.
PR 11-APR-1995; 95JP-0109128.
XX
PA (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
XX
PI Kubota M, Mitsuzumi H, Sugimoto T;
XX
DR WPI; 1996-106284/12.
DR N-PSDB; AAT16899.
XX
PT Recombinant thermostable enzyme from Sulfolobus acidocaldarius, -
PT releases trehalose from non-reducing saccharide at temps. exceeding
PT 55 degrees Centigrade

XX
PS Claim 2; Page 53-54; 74pp; English.
XX
CC A thermostable enzyme (AAR92755) of Sulfolobus acidocaldarius ATCC
CC 33909 releases trehalose from non-reducing saccharides having a
CC trehalose structure as an end unit and a degree of polymerisation of
CC at least 3. It has a mol.wt. of 54-64 kDa (SDS-PAGE), a pI of
CC 5.6-6.6 and is substantially not inactivated when incubated in aq.
CC solution (pH 7.0) at 85 deg for 60 min. Recombinant enzyme is
CC obtd. by expression of an isolated DNA fragment (AAT16899) in
CC host cells, pref. Escherichia coli, using e.g. vector pBluescript
CC II SK(+). The trehelose is useful as a sweetener.
XX
SQ Sequence 556 AA;

Query Match 100.0%; Score 2954; DB 17; Length 556;
Best Local Similarity 100.0%; Pred. No. 7.5e-267;
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MFSFGGNIEKNKGIFKLWAPYVNSVKLKLSKKLIPMEKNDEGFFEVEIDDIEENLTYSYI 60
Db 1 mfsfggnieknkgifklwapyvnsvklklskklipmekndegffeveiddieenltysyi 60

Qy 61 IEDKREIPDPASRYQPLGVHDKSQLIRTDYQILDLGKVKIEDLIIYELHVGTFSQEGNFK 120
Db 61 iedkreipdpasryqplgvhdksqlirtdyqildlgkvkiedliiyelhvgtfsqegnfk 120

Qy 121 GVIEKLDYLKDLGITGIELMPVAQFPGNRDWGYDGVFLYAVQNTYGGPWELAKLVNEAHK 180
Db 121 gviekldylkdlgitgielmpvaqfpgnrdwgydgvflyavqntyggpwelaklvneahk 180

Qy 181 RGIAVILDVVYNHIGPEGNYLLGLGPYFSDRYKTPWGLTFNFDDRGCDQVRKFILENVEY 240
Db 181 rgiavildvvynhigpegnyllglgpyfsdryktpwgltfnfddrgcdqvrkfilenvey 240

Qy 241 WFKTFKIDGLRLDAVHAIFDNSPKHILQEIAEKAHQLGKFVIAESDLNDPKIVKDDCGYK 300
Db 241 wfktfkidglrldavhaifdnspkhilqeiaekahqlgkfviaesdlndpkivkddcgyk 300

Qy 301 IDAQWVDDFHHAVHAFITKEKDYYYQDFGRIEDIEKTFKDVFVYDGKYSRYRGRTHGAPV 360
Db 301 idaqwvddfhhavhafitkekdyyyqdfgriediektfkdvfvydgkysryrgrthgapv 360

Qy 361 GDLPPRKFVVFIQNHDQVGNRGNGERLSILTDKTTYLMAATLYILSPYIPLIFMGEEYYE 420
Db 361 gdlpprkfvvfiqnhdqvgnrgngerlsiltdkttylmaatlyilspyiplifmgeeyye 420

Qy 421 TNPFFFFSDFSDPVLIKGVREGRLKENNQMIDPQSEEAFLKSKLSWKIDEEVLDYYKQLI 480
Db 421 tnpffffsdfsdpvlikgvregrlkennqmidpqseeaflksklswkideevldyykqli 480

Qy 481 NIRKRYNNCKRVKEVRREGNCITLIMEKIGIIASFDDIVINSKITGNLLIGIGFPKKLKK 540
Db 481 nirkrynnckrvkevrregncitlimekigiiasfddivinskitgnlligigfpkklkk 540

Qy 541 DELIKVNRGVGVYQLE 556
Db 541 delikvnrgvgvyqle 556

RESULT 2
AAR90620
ID AAR90620 standard; Protein; 556 AA.
XX
AC AAR90620;
XX
DT 29-JUN-1996 (first entry)
XX
DE Sulfolobus acidocaldarius amylase for alpha, alpha-trehalose prodn.
XX
KW transferase; amylase; Sulfolobus; production; alpha, alpha-trehalose;
KW malto-oligosaccharide; hydrolysis.
XX
OS Sulfolobus acidocaldarius.
XX
PN WO9534642-A.
XX
PD 21-DEC-1995.
XX
PF 14-JUN-1995; 95WO-JP01189.
XX
PR 21-APR-1995; 95JP-0120673.
PR 15-JUN-1994; 94JP-0133354.
PR 18-AUG-1994; 94JP-0194223.
PR 31-OCT-1994; 94JP-0290394.
PR 21-NOV-1994; 94JP-0286917.
PR 21-NOV-1994; 94JP-0311185.
XX
PA (KIRI) KIRIN BEER KK.
XX
DR WPI; 1996-049671/05.
DR N-PSDB; AAT12326.
XX
PT Sulfolobus spp. derived transferase and amylase - for production of
PT alpha, alpha-trehalose from malto-oligosaccharide(s)
XX
PS Claim 112; Page 244-250; 357pp; Japanese.
XX
CC The amylase is derived from Sulfolobus acidocaldarius. The amylase acts
CC on a saccharide having at least three sugar units, which are pref.
CC glucose units at the reducing end (the linkage between the first and
CC second glucose units is alpha-1, alpha-1, while the linkage between the
CC second and third glucose units is alpha-1,4), to hydrolyse alpha-1,4
CC linkages within the sugar chain, yielding alpha, alpha-trehalose and also
CC mono- and disaccharide hydrolysis products. The amylase has a mol. wt. of
CC 61 to 64 kDa. It is characterised by working at pH 4.5-5.5 and at 60-85
CC deg.C. It has an isoelectric point of 4.3-5.4 and retains at least
CC 100 percent activity after 6 hrs. at 80 deg.C. It is completely inhibited
CC by 5 mM copper sulphate. Use of a transferase and the amylase in
CC succession on suitable substrates such a malto-oligosaccharides, is
CC useful for the production of alpha, alpha-trehalose.
XX
SQ Sequence 556 AA;

Query Match 100.0%; Score 2954; DB 17; Length 556;
Best Local Similarity 100.0%; Pred. No. 7.5e-267;
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MFSFGGNIEKNKGIFKLWAPYVNSVKLKLSKKLIPMEKNDEGFFEVEIDDIEENLTYSYI 60
Db 1 mfsfggnieknkgifklwapyvnsvklklskklipmekndegffeveiddieenltysyi 60

Qy 61 IEDKREIPDPASRYQPLGVHDKSQLIRTDYQILDLGKVKIEDLIIYELHVGTFSQEGNFK 120
Db 61 iedkreipdpasryqplgvhdksqlirtdyqildlgkvkiedliiyelhvgtfsqegnfk 120

Qy 121 GVIEKLDYLKDLGITGIELMPVAQFPGNRDWGYDGVFLYAVQNTYGGPWELAKLVNEAHK 180
Db 121 gviekldylkdlgitgielmpvaqfpgnrdwgydgvflyavqntyggpwelaklvneahk 180

Qy 181 RGIAVILDVVYNHIGPEGNYLLGLGPYFSDRYKTPWGLTFNFDDRGCDQVRKFILENVEY 240
Db 181 rgiavildvvynhigpegnyllglgpyfsdryktpwgltfnfddrgcdqvrkfilenvey 240

Qy 241 WFKTFKIDGLRLDAVHAIFDNSPKHILQEIAEKAHQLGKFVIAESDLNDPKIVKDDCGYK 300
Db 241 wfktfkidglrldavhaifdnspkhilqeiaekahqlgkfviaesdlndpkivkddcgyk 300

Qy 301 IDAQWVDDFHHAVHAFITKEKDYYYQDFGRIEDIEKTFKDVFVYDGKYSRYRGRTHGAPV 360
Db 301 idaqwvddfhhavhafitkekdyyyqdfgriediektfkdvfvydgkysryrgrthgapv 360

Qy 361 GDLPPRKFVVFIQNHDQVGNRGNGERLSILTDKTTYLMAATLYILSPYIPLIFMGEEYYE 420

```
Db    361 gdlpprkfvvfiqnhdqvgnrgngerlsiltdkttylmaatlyilspyiplifmgeeyye 420

Qy    421 TNPFFFFSDFSDPVLIKGVREGRLKENNQMIDPQSEEAFLKSKLSWKIDEEVLDYYKQLI 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 tnpfffsdfsdpvlikgvregrlkennqmidpqseeaflksklswkideevldyykqli 480

Qy    481 NIRKRYNNCKRVKEVRREGNCITLIMEKIGIIASFDDIVINSKITGNLLIGIGFPKKLKK 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 nirkrynnckrvkevrregncitlimekigiiasfddivinskitgnlligigfpkklkk 540

Qy    541 DELIKVNRGVGVYQLE 556
          ||||||||||||||||
Db    541 delikvnrgvgvyqle 556


RESULT   3
AAR90619
ID   AAR90619 standard; Protein; 559 AA.
XX
AC   AAR90619;
XX
DT   29-JUN-1996  (first entry)
XX
DE   Sulfolobus solfataricus amylase for alpha, alpha-trehalose prodn.
XX
KW   transferase; amylase; Sulfolobus; production; alpha, alpha-trehalose;
KW   malto-oligosaccharide; hydrolysis.
XX
OS   Sulfolobus solfataricus.
XX
PN   WO9534642-A.
XX
PD   21-DEC-1995.
XX
PF   14-JUN-1995;   95WO-JP01189.
XX
PR   21-APR-1995;   95JP-0120673.
PR   15-JUN-1994;   94JP-0133354.
PR   18-AUG-1994;   94JP-0194223.
PR   31-OCT-1994;   94JP-0290394.
PR   21-NOV-1994;   94JP-0286917.
PR   21-NOV-1994;   94JP-0311185.
XX
PA   (KIRI ) KIRIN BEER KK.
XX
DR   WPI; 1996-049671/05.
DR   N-PSDB; AAT12325.
XX
PT   Sulfolobus spp. derived transferase and amylase - for production of
PT   alpha, alpha-trehalose from malto-oligosaccharide(s)
XX
PS   Claim 108; Page 235-240; 357pp; Japanese.
XX
CC   The amylase is derived from Sulfolobus solfataricus.  The amylase acts
CC   on a saccharide having at least three sugar units, which are pref.
CC   glucose units at the reducing end (the linkage between the first and
CC   second glucose units is alpha-1, alpha-1, while the linkage between the
CC   second and third glucose units is alpha-1,4), to hydrolyse alpha-1,4
CC   linkages within the sugar chain, yielding alpha, alpha-trehalose and also
CC   mono- and disaccharide hydrolysis products. The amylase has a mol. wt. of
CC   61 to 64 kDa.  It is characterised by working at pH 4.5-5.5 and at 60-85
CC   deg.C.  It has an isoelectric point of 4.3-5.4 and retains at least
CC   100 percent activity after 6 hrs. at 80 deg.C. It is completely inhibited
CC   by 5 mM copper sulphate. Use of a transferase and the amylase in
CC   succession on suitable substrates such a malto-oligosaccharides, is
CC   useful for the production of alpha, alpha-trehalose.
XX
SQ   Sequence   559 AA;


  Query Match             59.2%;  Score 1748.5;  DB 17;  Length 559;
  Best Local Similarity   58.9%;  Pred. No. 2.9e-154;
  Matches  329;  Conservative   87;  Mismatches  136;  Indels    7;  Gaps    5;

Qy      3 SFGGNIEKNKGIFKLWAPYVNSVKLK-LSKKLIPMEKNDEGFFEVEIDDIEENLTYSYII 61
          :|   |: |: || |||||  ||||| | | |  ||::::|:| : :::::    | |::
Db      2 tfaykidgneviftlwapyqksvklkvlekglyemerdekgyftitlnnvkvrdrykyvl 61

Qy     62 EDKREIPDPASRYQPLGVHDKSQLIRTDYQILDLGKVKIEDLIIYELHVGTFSQEGNFKG 121
          :|  |||||||||| ||| ||:|:   :  :   :| |||||||:|||||: || |:|
Db     62 ddaseipdpasryqpegvhgpsqiiqeskefnnetflkkedliiyeihvgtftpegtfeg 121

Qy    122 VIEKLDYLKDLGITGIELMPVAQFPGNRDWGYDGVFLYAVQNTYGGPWELAKLVNEAHKR 181
          || |||||||||| ||:||:|||| |||||||:||||||:|||    |||:||||:
Db    122 virkldylkdlgitaieimpiaqfpgkrdwgydgvylyavqnsyggpegfrklvdeahkk 181

Qy    182 GIAVILDVVYNHIGPEGNYLLGLGPYFSDRYKTPWGLTFNFDDRGCDQVRKFILENVEYW 241
          |: |||||||||:||||||:: |||||| :|||||||||||||   |:|||||||||||
Db    182 glgvildvvynhvgpegnymvklgpyfsqkyktpwgltfnfddaesdevrkfilenveyw 241

Qy    242 FKTFKIDGLRLDAVHAIFDNSPKHILQEIAEKAHQLGKFVIAESDLNDPKIV--KDDCGY 299
          | : :|| |||||||| | ||||||:|||:  |:  : ||||||||||::|  |: |||
Db    242 ikeynvdgfrldavhaiidtspkhileeiadvvhkynriviaesdlndprvvnpkekcgy 301

Qy    300 KIDAQWVDDFHHAVHAFITKEKDYYYQDFGRIEDIEKTFKDVFVYDGKYSRYRGRTHGAP 359
          ||||||||||::||::| |:  || ||| ::|| |::||||||||||| :| :||| |
Db    302 nidaqwvddfhhsihayltgerqgyytdfgnlddivksykdvfvydgkysnfrrkthgep 361

Qy    360 VGDLPPRKFVVFIQNHDQVGNRGNGERLSILTDKTTYLMAATLYILSPYIPLIFMGEEYY 419
          ||:|    |||:|||||||||| |||:  | |: :| :|| ||:||||||:|||||||
Db    362 vgeldgcnfvvyiqnhdqvgnrgkgeriiklvdresykiaaalyllspyipmifmgeeyg 421

Qy    420 ETNPFFFFSDFSDPVLIKGVREGRLKENNQMIDPQSEEAFLKSKLSWKIDEEVLDYYKQL 479
          | |||:|||||||  ||:|||||| ||| |  ||| |  |  ||||||||||:  :|| |
Db    422 eenpfyffsdfsdskliqgvregrkkengqdtdpqdestfnasklswkideeifsfykil 481

Qy    480 INIRKRYN-NCKRVKEVRREGNCITLIMEKIGIIASFDDIVINSKITGNLLIGI--GFPK 536
          | :||  :  | |   |    | : :   :   :  |    |  | :| ||:     ||:
Db    482 ikmrkelsiacdrrvnvvngenwliikgreyfslyvfskssievkysgtlllssnnsfpq 541

Qy    537 KLKKDELIKVNRGVGVYQL 555
           ::: :  : ::|  :|:|
Db    542 hieegk-yefdkgfalykl 559


RESULT   4
AAG92072
ID   AAG92072 standard; Protein; 595 AA.
XX
AC   AAG92072;
XX
DT   26-SEP-2001  (first entry)
XX
DE   C glutamicum protein fragment SEQ ID NO: 5826.
XX
KW   Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
KW   organic acid synthesis.
XX
OS   Corynebacterium glutamicum.
XX
PN   EP1108790-A2.
XX
PD   20-JUN-2001.
XX
PF   18-DEC-2000; 2000EP-0127688.
XX
PR   16-DEC-1999;   99JP-0377484.
PR   07-APR-2000; 2000JP-0159162.
PR   03-AUG-2000; 2000JP-0280988.
XX
PA   (KYOW ) KYOWA HAKKO KOGYO KK.
XX
PI   Nakagawa S,  Mizoguchi H,  Ando S,  Hayashi M,  Ochiai K,  Yokoi H;
PI   Tateishi N,  Senoh A,  Ikeda M,  Ozaki A;
XX
DR   WPI; 2001-376931/40.
```

```
DR   N-PSDB; AAH67291.
XX
PT   Novel polynucleotides derived from Coryneform bacteria, for identifying
PT   mutation point of a gene, measuring expression of a gene, analysing
PT   expression profile or pattern of a gene and identifying homologous gene
PT   -
XX
PS   Claim 17; SEQ ID NO: 5826; 246pp + Sequence Listing; English.
XX
CC   The present invention provides a number of nucleotide and protein
CC   sequences from the Coryneform bacterium Corynebacterium glutamicum. These
CC   are useful for identifying the mutation point of a gene derived from a
CC   mutant of coryneform bacterium, measuring expression amount and
CC   analysing the expression profile or expression pattern of a gene derived
CC   from Coryneform bacterium, and identifying a homologue of a gene derived
CC   from coryneform bacterium. Coryneform bacteria are useful for producing
CC   amino acids, nucleic acids, vitamins, saccharides and organic acids,
CC   particularly L-lysine. The present sequence is a protein described
CC   in the exemplification of the invention.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from the
CC   European Patent Office.
XX
SQ   Sequence   595 AA;


  Query Match             35.3%;  Score 1042.5;  DB 22;  Length 595;
  Best Local Similarity   40.8%;  Pred. No. 2.7e-88;
  Matches  225;  Conservative   93;  Mismatches  199;  Indels   35;  Gaps   13;

Qy      15 FKLWAPYVNSVKLKLSKKLIPMEKNDEGFFEVEIDDIEENLTYSYIIED----KREIPDP 70
           | :|||  : | | |: : :|| | :  ::  ||   :    | : : |     : :|||
Db      28 fsvwaplphdvhlilngetlpmhktegswwraeiap-kagdrygfslfdgsswsktlpdp 86

Qy      71 ASRYQPLGVHDKSQLIRTDYQILD---LGKVKIEDLIIYELHVGTFSQEGNFKGVIEKLD 127
             |  || |||  |::    |   |    |:: :   ::||||||||||::| |:||::||
Db      87 rstsqpdgvhglsevsddsylwgdqqwtgri-lpgsvlyelhvgtfsedgtfegvvdklp 145

Qy     128 YLKDLGITGIELMPVAQFPGNRDWGYDGVFLYAVQNTYGGPWELAKLVNEAHKRGIAVIL 187
           ||:|||:| |||:||  | |||:||||||  :||   ||||  | ||:: :|: |||| |
Db     146 ylrdlgvtaiellpvqpfggnrnwgydgvlwhavhagyggpaglkklidashqagiavyl 205

Qy     188 DVVYNHIGPEGNYLLGLGPYFSDRYKTPWGLTFNFDDRGCDQVRKFILENVEYWFKTFKI 247
           ||||||  ||:|||    ||| |    |·||   | :    |:|| :||:    ||: | :
Db     206 dvvynhfgpdgnyngqfgpytsgg-stgwgdvvninghdsdevrnyildaarqwfedfhv 264

Qy     248 DGLRLDAVHAIFDNSPKHILQEIAEKAHQLG------KFVIAESDLNDPKIV--KDDCGY 299
           |||||||||:: |     :| ::   |  :       : :||||:||||| |  ::  |:
Db     265 dglrldavhsldrgaysllaqltmvaedvsaqtgiprsliaeselndpkfvtsreaggf 324

Qy     300 KIDAQWVDDFHHAVHAFITKEKDYYYQDFGRIEDIEKTFKDVFVYDGKYSRYRGRTHGAP 359
           :|||||||  |||:|| :: |:: || ||| :: : || ::|| : | || |||| || |
Db     325 gldaqwvddihhalhalvsgerngyysdfgsvdtlaktlrevfehtgnystyrgrnhgrp 384

Qy     360 VGD--LPPRKFVVFIQNHDQVGNRGNGERLS-ILTDKTTYLMAATLYILSPYIPLIFMGE 416
           |     |  :|| :   ||| |||  |:| |  || :   | || :|  ||| |::||||
Db     385 vhpditpasrfvtytthdqtgnraigdrpsttltpeqqvlkaaiiy-sspytpmlfmge 443

Qy     417 EYYETNPFFFFSDFSDPVLIKGVREGRLKE-------NNQMIDPQSEEAFLKSKLSWKID 469
           |:  | || ||   :|| | :   ||| :|        : :  |: |  |  ||| |:
Db     444 efgattpfaffcshtdpelnrltsegrkrefarlgwnaddipspelestftsskldweft 503

Qy     470 EE---VLDYYKQLINIRKRYNNCK---RVKEVRREGNCITLIMEKIGIIASFDDIVINSK 523
             |   : | ||||:::|      :      ||    | :::   :  |:|:| |  |
Db     504 aeqrrindaykqllhlrhtlgfsqpnlltlevehgenwlsmangrgrilanfsddtitvp 563

Qy     524 ITGNLLIGIGFP 535
           : | |:     |
Db     564 lggeliysftsp 575


RESULT   5
AAB80206
```

```
ID   AAB80206 standard; Protein; 610 AA.
XX
AC   AAB80206;
XX
DT   30-APR-2001  (first entry)
XX
DE   Corynebacterium glutamicum MP protein sequence SEQ ID NO:1146.
XX
KW   Corynebacterium glutamicum; metabolic pathway protein; MP protein;
KW   fine chemical production; microorganism; organic acid; nucleoside;
KW   nonproteinogenic amino acid; purine base; pyrimidine base; nucleotide;
KW   lipid; saturated fatty acid; unsaturated fatty acid; diol; vitamin;
KW   carbohydrate; aromatic compound; cofactor; polyketide; enzyme.
XX
OS   Corynebacterium glutamicum.
XX
PN   WO200100843-A2.
XX
PD   04-JAN-2001.
XX
PF   23-JUN-2000; 2000WO-IB00923.
XX
PR   25-JUN-1999;   99US-0141031.
PR   01-JUL-1999;   99DE-1030476.
PR   02-JUL-1999;   99US-0142101.
PR   08-JUL-1999;   99DE-1031415.
PR   08-JUL-1999;   99DE-1031418.
PR   08-JUL-1999;   99DE-1031419.
PR   08-JUL-1999;   99DE-1031420.
PR   08-JUL-1999;   99DE-1031424.
PR   08-JUL-1999;   99DE-1031428.
PR   08-JUL-1999;   99DE-1031434.
PR   08-JUL-1999;   99DE-1031435.
PR   08-JUL-1999;   99DE-1031443.
PR   08-JUL-1999;   99DE-1031453.
PR   08-JUL-1999;   99DE-1031457.
PR   08-JUL-1999;   99DE-1031465.
PR   08-JUL-1999;   99DE-1031478.
PR   08-JUL-1999;   99DE-1031510.
PR   08-JUL-1999;   99DE-1031541.
PR   08-JUL-1999;   99DE-1031573.
PR   08-JUL-1999;   99DE-1031592.
PR   08-JUL-1999;   99DE-1031632.
PR   08-JUL-1999;   99DE-1031634.
PR   08-JUL-1999;   99DE-1031636.
PR   09-JUL-1999;   99DE-1032125.
PR   09-JUL-1999;   99DE-1032126.
PR   09-JUL-1999;   99DE-1032130.
PR   09-JUL-1999;   99DE-1032186.
PR   09-JUL-1999;   99DE-1032206.
PR   09-JUL-1999;   99DE-1032227.
PR   09-JUL-1999;   99DE-1032228.
PR   09-JUL-1999;   99DE-1032229.
PR   09-JUL-1999;   99DE-1032230.
PR   14-JUL-1999;   99DE-1032922.
PR   14-JUL-1999;   99DE-1032926.
PR   14-JUL-1999;   99DE-1032928.
PR   14-JUL-1999;   99DE-1033004.
PR   14-JUL-1999;   99DE-1033005.
PR   14-JUL-1999;   99DE-1033006.
PR   12-AUG-1999;   99US-0148613.
PR   27-AUG-1999;   99DE-1040764.
PR   27-AUG-1999;   99DE-1040765.
PR   27-AUG-1999;   99DE-1040766.
PR   27-AUG-1999;   99DE-1040832.
PR   31-AUG-1999;   99DE-1041378.
PR   31-AUG-1999;   99DE-1041379.
PR   31-AUG-1999;   99DE-1041380.
PR   31-AUG-1999;   99DE-1041394.
PR   31-AUG-1999;   99DE-1041396.
PR   03-SEP-1999;   99DE-1042076.
PR   03-SEP-1999;   99DE-1042077.
PR   03-SEP-1999;   99DE-1042079.
```

PR 03-SEP-1999; 99DE-1042086.
PR 03-SEP-1999; 99DE-1042087.
PR 03-SEP-1999; 99DE-1042088.
PR 03-SEP-1999; 99DE-1042095.
PR 03-SEP-1999; 99DE-1042124.
PR 03-SEP-1999; 99DE-1042129.
PR 09-MAR-2000; 2000US-0187970.
XX
PA (BADI) BASF AG.
XX
PI Pompejus M, Kroeger B, Schroeder H, Zelder O, Haberhauer G;
XX
DR WPI; 2001-137957/14.
DR N-PSDB; AAF72325.
XX
PT Nucleic acids from Corynebacterium glutamicum encoding metabolic
PT pathway proteins, useful for producing fine chemicals in
PT microorganisms, including organic acids, nonproteinogenic amino acids,
PT and purine and pyrimidine bases -
XX
PS Claim 20; Page 1722-1725; 1737pp; English.
XX
CC AAF71753 to AAF72330 encode the Corynebacterium glutamicum metabolic
CC pathway (MP) proteins given in AAB79634 to AAB80211. The C. glutamicum
CC MP nucleic acids are useful for the production of fine chemicals
CC in microorganisms, including organic acids, nonproteinogenic amino
CC acids, purine and pyrimidine bases, nucleosides, nucleotides, lipids,
CC saturated and unsaturated fatty acids, diols, carbohydrates, aromatic
CC compounds, vitamins, cofactors, polyketides and enzymes.
XX
SQ Sequence 610 AA;

Query Match 35.3%; Score 1042.5; DB 22; Length 610;
Best Local Similarity 40.8%; Pred. No. 2.8e-88;
Matches 225; Conservative 93; Mismatches 199; Indels 35; Gaps 13;

```
Qy     15 FKLWAPYVNSVKLKLSKKLIPMEKNDEGFFEVEIDDIEENLTYSYIIED----KREIPDP 70
Db     43 fsvwaplphdvhliingetlpmhktegswwraeiap-kagdrygfslfdgsswsktlpdp 101

Qy     71 ASRYQPLGVHDKSQLIRTDYQILD---LGKVKIEDLIIYELHVGTFSQEGNFKGVIEKLD 127
Db    102 rstsqpdgvhglsevsddsylwgdqqwtgri-lpgsvlyelhvgtfsedgtfegvvdklp 160

Qy    128 YLKDLGITGIELMPVAQFPGNRDWGYDGVFLYAVQNTYGGPWELAKLVNEAHKRGIAVIL 187
Db    161 ylrdlgvtaiellpvqpfggnrnwgydgvlwhavhagyggpaglkklidashqagiavyl 220

Qy    188 DVVYNHIGPEGNYLLGLGPYFSDRYKTPWGLTFNFDDRGCDQVRKFILENVEYWFKTFKI 247
Db    221 dvvynhfgpdgnyngqfgpytsgg-stgwgdvvninghdsdevrnyildaarqwfedfhv 279

Qy    248 DGLRLDAVHAIFDNSPKHILQEIAEKAHQLG------KFVIAESDLNDPKIV--KDDCGY 299
Db    280 dglrldavhsldrgaysllaqltmvaedvsaqtgiprsliaeselndpkfvtsreaggf 339

Qy    300 KIDAQWVDDFHHAVHAFITKEKDYYYQDFGRIEDIEKTFKDVFVYDGKYSRYRGRTHGAP 359
Db    340 gldaqwvddihhalhalvsgerngyysdfgsvdtlaktlrevfehtgnystyrgrnhgrp 399

Qy    360 VGD--LPPRKFVVFIQNHDQVGNRGNGERLS-ILTDKTTYLMAATLYILSPYIPLIFMGE 416
Db    400 vhpditpasrfvtyttthdqtgnraigdrpsttltpeqqvlkaaiiy-sspytpmlfmge 458

Qy    417 EYYETNPFFFFSDFSDPVLIKGVREGRLKE-------NNQMIDPQSEEAFLKSKLSWKID 469
Db    459 efgattpfaffcshtdpelnrltsegrkrefarlgwnaddipspelestftsskldweft 518

Qy    470 EE---VLDYYKQLINIRKRYNNCK---RVKEVRREGNCITLIMEKIGIIASFDDIVINSK 523
Db    519 aeqrrindaykqllhlrhtlgfsqpnlltlevehgenwlsmangrgrilanfsddtitvp 578
```

Qy 524 ITGNLLIGIGFP 535
Db 579 lggeliysftsp 590

RESULT 6
AAR80289
ID AAR80289 standard; Protein; 589 AA.
XX
AC AAR80289;
XX
DT 19-JAN-1996 (first entry)
XX
DE Trehalose releasing enzyme.
XX
KW Enzyme; trehalose; alpha-glucosyltrehalose; alpha-maltosyltrehalose;
KW alpha-maltotriosyltrehalose; alpha-maltotetrasyltrehalose;
KW maltopentaosyltrehalose; sweetener; taste-improving agent;
KW stabiliser; filler; excipient; adjuvant; foodstuffs; cosmetics;
KW pharmaceuticals.
XX
OS Rhizobium sp. M11.
XX
PN EP671470-A2.
XX
PD 13-SEP-1995.
XX
PF 07-MAR-1995; 95EP-0301474.
XX
PR 07-MAR-1994; 94JP-0059840.
PR 07-MAR-1994; 94JP-0059834.
XX
PA (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
XX
PI Hattori K, Kubota M, Sugimoto T, Tsusaki K;
XX
DR WPI; 1995-312772/41.
DR N-PSDB; AAQ98669.
XX
PT DNA encoding a trehalose releasing enzyme - which releases trehalose
PT from a non-reducing saccharide having a trehalose structure as an
PT end unit.
XX
PS Claim 3; Page 21-22; 45pp; English.
XX
CC This enzyme can be used for the preparation of trehalose with high
CC yields and efficiency from non-reducing saccharides such as
CC alpha-glucosyltrehalose, alpha-maltosyltrehalose,
CC alpha-maltotriosyltrehalose, alpha-maltotetrasyltrehalose and
CC maltopentaosyltrehalose. The trehalose can be used as a sweetener,
CC taste-improving agent, quality-improving agent, stabiliser, filler,
CC excipient or adjuvant in food products cosmetics and pharmaceuticals.
XX
SQ Sequence 589 AA;

Query Match 32.8%; Score 969; DB 16; Length 589;
Best Local Similarity 39.5%; Pred. No. 2e-81;
Matches 215; Conservative 74; Mismatches 195; Indels 60; Gaps 12;

```
Qy     13 GIFKLWAPYVNSVKLKLSKKLIPMEKN------DEGFFEVEIDDIEENLTYSYIIE-DKR 65
Db      8 grfdiwapeagtvtllaggeryemgrrpgngpadegwwtaadaptgadvdygylldgdei 67

Qy     66 EIPDPASRYQPLGVHDKSQLIRTDYQILDLGKVKIEDL----------IIYELHVGTFSQ 115
Db     68 plpdprtrrqpegvh---alsrt----fdpgahrwqdagwqgrelqgsviyelhigtftp 120

Qy    116 EGNFKGVIEKLDYLKDLGITGIELMPVAQFPGNRDWGYDGVFLYAVQNTYGGPWELAKLV 175
Db    121 egtldaaagkldylaglgidfiellpvnafngthnwgydgvqwfavhegyggpaayqrfv 180

Qy    176 NEAHKRGIAVILDVVYNHIGPEGNYLLGLGPYFSDRYKTPWGLTFNFDDRGCDQVRKFIL 235
```

```
          : ||  |: || ||||||:|| ||||   |||        || : | |  | | ||::||
Db    181 daahaaglgviqdvvynhlgpsgnylprygpylkhgegntwgdsvnldgpgsdhvrqyil 240

Qy    236 ENVEYWFKTFKIDGLRLDAVHAIFDNSPKHILQEIAEKAHQLGK------FVIAESDLND 289
          :||  | : :::|||||||||||: |    |||:|   |  |         :|||||||:
Db    241 dnvamwlrdyrvdglrldavhalkderavhileefgaladalsseggrpltliaesdlnn 300

Qy    290 PKIV--KDDCGYKIDAQWVDDFHHAVHAFITKEKDYYYQDFGRIEDIEKTFKDVFVYDGK 347
          |:::  :|  || :  || ||||||||  :: |   || ||  :  : |  :| | :||
Db    301 prllyprdvngyglagqwsddfhhavhvnvsgettgyysdfdslgalakvlrdgffhdgs 360

Qy    348 YSRYRGRTHGAPV--GDLPPRKFVVFIQNHDQVGNRGNGERLSILTDKTTYLMAATLYIL 405
          || :||| || |:    : |   ||  |||||:|||  |:|||      :  :|| | :
Db    361 yssfrgrchgrpinfsavhpaalvvcsqnhdqignratgdrlsqslpygslalaavltlt 420

Qy    406 SPYIPLIFMGEEYYETNPFFFFSDFSDPVLIKGVREGRLKENNQM-------IDPQSEEA 458
           |: |::|||||  | |: ||:   :| | |   |||::|  :|        |||  |
Db    421 gpftpmlfmgeeygattpwqfftshpepelgkataegrirefermgwdpavvpdpqdpet 480

Qy    459 FLKSKLSWK-----IDEEVLDYYKQLINIRKRYNNCKRVKEVRREGNCITLIMEKIGIIA 513
          | :||| |          :|: |: || :|      :   |: | |   | :
Db    481 ftrskldwaeasagdharllelyrslitlr------rstpelarlgfadtav-------- 526

Qy    514 SFDD 517
           |||
Db    527 efdd 530


RESULT   7
AAR77470
ID   AAR77470 standard; Protein; 596 AA.
XX
AC   AAR77470;
XX
DT   25-JAN-1996  (first entry)
XX
DE   Trehalose releasing enzyme.
XX
KW   Enzyme; trehalose; alpha-glucosyltrehalose; alpha-maltosyltrehalose;
KW   alpha-maltotriosyltrehalose; alpha-maltotetrasyltrehalose;
KW   maltopentaosyltrehalose; sweetener; taste-improving agent;
KW   stabiliser; filler; excipient; adjuvant; foodstuffs; cosmetics;
KW   pharmaceuticals.
XX
OS   Rhizobium sp. M11.
XX
PN   EP671470-A2.
XX
PD   13-SEP-1995.
XX
PF   07-MAR-1995;   95EP-0301474.
XX
PR   07-MAR-1994;   94JP-0059840.
PR   07-MAR-1994;   94JP-0059834.
XX
PA   (HAYB ) HAYASHIBARA SEIBUTSU KAGAKU.
XX
PI   Hattori K,  Kubota M,  Sugimoto T,  Tsusaki K;
XX
DR   WPI; 1995-312772/41.
DR   N-PSDB; AAQ98671.
XX
PT   DNA encoding a trehalose releasing enzyme - which releases trehalose
PT   from a non-reducing saccharide having a trehalose structure as an
PT   end unit.
XX
PS   Claim 3; Page 27-29; 45pp; English.
XX
CC   This enzyme can be used for the preparation of trehalose with high
CC   yields and efficiency from non-reducing saccharides such as
CC   alpha-glucosyltrehalose, alpha-maltosyltrehalose,
CC   alpha-maltotriosyltrehalose, alpha-maltotetrasyltrehalose and
CC   maltopentaosyltrehalose.  The trehalose can be used as a sweetener,
CC   taste-improving agent, quality-improving agent, stabiliser, filler,
CC   excipient or adjuvant in food products cosmetics and pharmaceuticals.
XX
SQ   Sequence   596 AA;


  Query Match             32.8%;  Score 969;  DB 16;  Length 596;
  Best Local Similarity   39.5%;  Pred. No. 2e-81;
  Matches  215;  Conservative   74;  Mismatches  195;  Indels   60;  Gaps   12;

Qy     13 GIFKLWAPYVNSVKLKLSKKLIPMEKN------DEGFFEVEIDDIEENLTYSYIIE-DKR 65
          | | :|||   :| |    :   | :       |||::          :: | |::: |:
Db     15 grfdiwapeagtvtllaggeryemgrrpgngpadegwwtaadaptgadvdygylldgdei 74

Qy     66 EIPDPASRYQPLGVHDKSQLIRTDYQILDLGKVKIEDL----------IIYELHVGTFSQ 115
           :||| :| || |||    | ||     | |  : :|           :|||||:|||:
Db     75 plpdprtrrqpegvh---alsrt----fdpgahrwqdagwqgrelqgsviyelhigtftp 127

Qy    116 EGNFKGVIEKLDYLKDLGITGIELMPVAQFPGNRDWGYDGVFLYAVQNTYGGPWELAKLV 175
          ||       |||||  |||  |||:||  | |  :||||||  :||   ||||    : |
Db    128 egtldaaagkldylaglgidfiellpvnafngthnwgydgvqwfavhegyggpaayqrfv 187

Qy    176 NEAHKRGIAVILDVVYNHIGPEGNYLLGLGPYFSDRYKTPWGLTFNFDDRGCDQVRKFIL 235
          : ||  |: || ||||||:|| ||||   |||        || : | |  | | ||::||
Db    188 daahaaglgviqdvvynhlgpsgnylprygpylkhgegntwgdsvnldgpgsdhvrqyil 247

Qy    236 ENVEYWFKTFKIDGLRLDAVHAIFDNSPKHILQEIAEKAHQLGK------FVIAESDLND 289
          :||  | : :::|||||||||||: |    |||:|    |  |         :|||||||:
Db    248 dnvamwlrdyrvdglrldavhalkderavhileefgaladalsseggrpltliaesdlnn 307

Qy    290 PKIV--KDDCGYKIDAQWVDDFHHAVHAFITKEKDYYYQDFGRIEDIEKTFKDVFVYDGK 347
          |:::  :|  || :  || ||||||||  :: |   || ||  :  : |  :| | :||
Db    308 prllyprdvngyglagqwsddfhhavhvnvsgettgyysdfdslgalakvlrdgffhdgs 367

Qy    348 YSRYRGRTHGAPV--GDLPPRKFVVFIQNHDQVGNRGNGERLSILTDKTTYLMAATLYIL 405
          || :||| || |:    : |   ||  |||||:|||  |:|||      :  :|| | :
Db    368 yssfrgrchgrpinfsavhpaalvvcsqnhdqignratgdrlsqslpygslalaavltlt 427

Qy    406 SPYIPLIFMGEEYYETNPFFFFSDFSDPVLIKGVREGRLKENNQM-------IDPQSEEA 458
           |: |::|||||  | |: ||:   :| | |   |||::|  :|        |||  |
Db    428 gpftpmlfmgeeygattpwqfftshpepelgkataegrirefermgwdpavvpdpqdpet 487

Qy    459 FLKSKLSWK-----IDEEVLDYYKQLINIRKRYNNCKRVKEVRREGNCITLIMEKIGIIA 513
          | :||| |          :|: |: || :|      :   |: | |   | :
Db    488 ftrskldwaeasagdharllelyrslitlr------rstpelarlgfadtav-------- 533

Qy    514 SFDD 517
           |||
Db    534 efdd 537


RESULT   8
AAR77471
ID   AAR77471 standard; Protein; 598 AA.
XX
AC   AAR77471;
XX
DT   25-JAN-1996  (first entry)
XX
DE   Trehalose releasing enzyme.
XX
KW   Enzyme; trehalose; alpha-glucosyltrehalose; alpha-maltosyltrehalose;
KW   alpha-maltotriosyltrehalose; alpha-maltotetrasyltrehalose;
KW   maltopentaosyltrehalose; sweetener; taste-improving agent;
KW   stabiliser; filler; excipient; adjuvant; foodstuffs; cosmetics;
KW   pharmaceuticals.
XX
OS   Arthrobacter sp. Q36.
XX
PN   EP671470-A2.
XX
PD   13-SEP-1995.
XX
```

```
PF   07-MAR-1995;   95EP-0301474.
XX
PR   07-MAR-1994;   94JP-0059840.
PR   07-MAR-1994;   94JP-0059834.
XX
PA   (HAYB ) HAYASHIBARA SEIBUTSU KAGAKU.
XX
PI   Hattori K,  Kubota M,  Sugimoto T,  Tsusaki K;
XX
DR   WPI; 1995-312772/41.
DR   N-PSDB; AAQ98672.
XX
PT   DNA encoding a trehalose releasing enzyme - which releases trehalose
PT   from a non-reducing saccharide having a trehalose structure as an
PT   end unit.
XX
PS   Claim 6; Page 27-29; 45pp; English.
XX
CC   This enzyme can be used for the preparation of trehalose with high
CC   yields and efficiency from non-reducing saccharides such as
CC   alpha-glucosyltrehalose, alpha-maltosyltrehalose,
CC   alpha-maltotriosyltrehalose, alpha-maltotetrasyltrehalose and
CC   maltopentaosyltrehalose.  The trehalose can be used as a sweetener,
CC   taste-improving agent, quality-improving agent, stabiliser, filler,
CC   excipient or adjuvant in food products cosmetics and pharmaceuticals.
XX
SQ   Sequence   598 AA;

  Query Match             32.5%;  Score 959;  DB 16;  Length 598;
  Best Local Similarity   40.2%;  Pred. No. 1.7e-80;
  Matches  202;  Conservative   81;  Mismatches  186;  Indels   34;  Gaps   10;

Qy      15 FKLWAPYVNSVKLKLSKKLIPMEK------NDEGFFEVEIDDIEENLTYSYIIE-DKREI 67
           : :|||   || |    :   |::       | |::       : |: | |::: |:  :
Db      19 ydvwapnaesvtllaggeryamqrraetgpedagwwtaagaptdgnvdygylldgdetpl 78

Qy      68 PDPASRYQPLGVHDKSQLI-RTDYQILD---LGKVKIEDLIIYELHVGTFSQEGNFKGVI 123
           ||| :| || |||  |:    : |   |    |: :::  :|||||:|||: ||  :
Db      79 pdprtrrqpdgvhalsrtfdpsayswqddawqgr-elqgaviyelhlgtftpegtleaaa 137

Qy     124 EKLDYLKDLGITGIELMPVAQFPGNRDWGYDGVFLYAVQNTYGGPWELAKLVNEAHKRGI 183
           |||||  ||:  |||:||  | |  :||||||  :||   ||||    : |: ||  |:
Db     138 gkldylaglgvdfiellpvnafngthnwgydgvqwfavheayggpeayqrfvdaahaagl 197

Qy     184 AVILDVVYNHIGPEGNYLLGLGPYFSDRYKTPWGLTFNFDDRGCDQVRKFILENVEYWFK 243
           || ||||||:|| ||||   |||        || : | |  | | ||::||:|:  | :
Db     198 gviqdvvynhlgpsgnylprfgpylkqgegntwgdsvnldgpgsdhvrryildnlamwlr 257

Qy     244 TFKIDGLRLDAVHAIFDNSPKHILQEIAEKAHQLGKFV------IAESDLNDPKIV--KD 295
           :::|||||||||||: |    |||::    | |:   |      |||||||:|:::  :|
Db     258 dyrvdglrldavhalkderavhiledfgaladqisaevgrpltliaesdlnnprllyprd 317

Qy     296 DCGYKIDAQWVDDFHHAVHAFITKEKDYYYQDFGRIEDIEKTFKDVFVYDGKYSRYRGRT 355
             || :: || ||||||||  :| |   || ||  :  : |  :| | :|| || :| |
Db     318 vngyglegqwsddfhhavhvnvtgettgyysdfdslaalakvlrdgffhdgsyssfrerh 377

Qy     356 HGAPV--GDLPPRKFVVFIQNHDQVGNRGNGERLSILTDKTTYLMAATLYILSPYIPLIF 413
           || |:    : |   ||  ||||:|||  |:|||      :  :|| | :  |: |::
Db     378 hgrpinfsavhpaalvvcsqnhdqignratgdrlsqtlpygslalaavltltgpftpmll 437

Qy     414 MGEEYYETNPFFFFSDFSDPVLIKGVREGRLKENNQM-------IDPQSEEAFLKSKLSW 466
           |||||  : |: ||:   :| | |   |||:|| :|        |||  | | :|||  |
Db     438 mgeeygastpwqfftshpepelgkataegrikefermgwdpavvpdpqdpetfrrskldw 497

Qy     467 KIDEE-----VLDYYKQLINIRK 484
               |     :|: |: |  :|:
Db     498 aeaaegdharllelyrsltalrr 520


RESULT   9
AAR80290
ID   AAR80290 standard; Protein; 597 AA.
XX
AC   AAR80290;
XX
DT   19-JAN-1996  (first entry)
XX
DE   Trehalose releasing enzyme.
XX
KW   Enzyme; trehalose; alpha-glucosyltrehalose; alpha-maltosyltrehalose;
KW   alpha-maltotriosyltrehalose; alpha-maltotetrasyltrehalose;
KW   maltopentaosyltrehalose; sweetener; taste-improving agent;
KW   stabiliser; filler; excipient; adjuvant; foodstuffs; cosmetics;
KW   pharmaceuticals.
XX
OS   Arthrobacter sp. Q36.
XX
PN   EP671470-A2.
XX
PD   13-SEP-1995.
XX
PF   07-MAR-1995;   95EP-0301474.
XX
PR   07-MAR-1994;   94JP-0059840.
PR   07-MAR-1994;   94JP-0059834.
XX
PA   (HAYB ) HAYASHIBARA SEIBUTSU KAGAKU.
XX
PI   Hattori K,  Kubota M,  Sugimoto T,  Tsusaki K;
XX
DR   WPI; 1995-312772/41.
DR   N-PSDB; AAQ98670.
XX
PT   DNA encoding a trehalose releasing enzyme - which releases trehalose
PT   from a non-reducing saccharide having a trehalose structure as an
PT   end unit.
XX
PS   Claim 3; Page 24-25; 45pp; English.
XX
CC   This enzyme can be used for the preparation of trehalose with high
CC   yields and efficiency from non-reducing saccharides such as
CC   alpha-glucosyltrehalose, alpha-maltosyltrehalose,
CC   alpha-maltotriosyltrehalose, alpha-maltotetrasyltrehalose and
CC   maltopentaosyltrehalose.  The trehalose can be used as a sweetener,
CC   taste-improving agent, quality-improving agent, stabiliser, filler,
CC   excipient or adjuvant in food products cosmetics and pharmaceuticals.
XX
SQ   Sequence   597 AA;

  Query Match             32.4%;  Score 958;  DB 16;  Length 597;
  Best Local Similarity   40.2%;  Pred. No. 2.1e-80;
  Matches  202;  Conservative   81;  Mismatches  186;  Indels   34;  Gaps   10;

Qy      15 FKLWAPYVNSVKLKLSKKLIPMEK------NDEGFFEVEIDDIEENLTYSYIIE-DKREI 67
           : :|||   || |    :   |::       | |::       : |: | |::: |:  :
Db      18 ydvwapnaesvtllaggeryamqrraetgpedagwwtaagaptdgnvdygylldgdetpl 77

Qy      68 PDPASRYQPLGVHDKSQLI-RTDYQILD---LGKVKIEDLIIYELHVGTFSQEGNFKGVI 123
           ||| :| || |||  |:    : |   |    |: :::  :|||||:|||: ||  :
Db      78 pdprtrrqpdgvhalsrtfdpsayswqddawqgr-elqgaviyelhlgtftpegtleaaa 136

Qy     124 EKLDYLKDLGITGIELMPVAQFPGNRDWGYDGVFLYAVQNTYGGPWELAKLVNEAHKRGI 183
           |||||  ||:  |||:||  | |  :||||||  :||   ||||    : |: ||  |:
Db     137 gkldylaglgvdfiellpvnafngthnwgydgvqwfavhedyggpeayqrfvdaahaagl 196

Qy     184 AVILDVVYNHIGPEGNYLLGLGPYFSDRYKTPWGLTFNFDDRGCDQVRKFILENVEYWFK 243
           || ||||||:|| ||||   |||        || : | |  | | ||::||:|:  | :
Db     197 gviqdvvynhlgpsgnylprfgpylkqgegntwgdsvnldgpgsdhvrryildnlamwlr 256

Qy     244 TFKIDGLRLDAVHAIFDNSPKHILQEIAEKAHQLGKFV------IAESDLNDPKIV--KD 295
           :::|||||||||||: |    |||::    | |:   |      |||||||:|:::  :|
Db     257 dyrvdglrldavhalkderavhiledfgaladqisaevgrpltliaesdlnnprllyprd 316

Qy     296 DCGYKIDAQWVDDFHHAVHAFITKEKDYYYQDFGRIEDIEKTFKDVFVYDGKYSRYRGRT 355
```

```
           || :: || |||||||| :| |   || ||  :  : |  :| | :|| || :| |
Db    317 vngyglegqwsddfhhavhvnvtgettgyysdfdslaalakvlrdgffhdgsyssfrerh 376

Qy    356 HGAPV--GDLPPRKFVVFIQNHDQVGNRGNGERLSILTDKTTYLMAATLYILSPYIPLIF 413
          || |:    : |   ||  |||||:|||  |:|||      :  :|| | :  |: |::
Db    377 hgrpinfsavhpaalvvcsqnhdqignratgdrlsqtlpygslalaavltltgpftpmll 436

Qy    414 MGEEYYETNPFFFFSDFSDPVLIKGVREGRLKENNQM-------IDPQSEEAFLKSKLSW 466
          |||||  : |: ||:   :| | |   |||:||  :|        |||  | | :||| |
Db    437 mgeeygastpwqfftshpepelgkataegrikefermgwdpavvpdpqdpetfrrskldw 496

Qy    467 KIDEE-----VLDYYKQLINIRK 484
            |     :|: |: |  :|:
Db    497 aeaaegdharllelyrsltalrr 519


RESULT 10
AAY85157
ID   AAY85157 standard; protein; 575 AA.
XX
AC   AAY85157;
XX
DT   23-JUN-2000  (first entry)
XX
DE   Trehalose-releasing enzyme amino acid sequence.
XX
KW   Non-reducing saccharide-forming enzyme; trehalose; disaccharide; food;
KW   high yield; trehalose releasing enzyme; moisture retention; cosmetic;
KW   production.
XX
OS   Arthrobacter sp. S34.
XX
PN   EP990704-A2.
XX
PD   05-APR-2000.
XX
PF   13-SEP-1999;   99EP-0307220.
XX
PR   11-SEP-1998;   98JP-0258394.
PR   11-DEC-1998;   98JP-0352252.
PR   26-JAN-1999;   99JP-0016931.
XX
PA   (HAYB ) HAYASHIBARA SEIBUTSU KAGAKU.
XX
PI   Yamamoto T,  Maruta K,  Kubota M,  Fukuda S,  Miyake T;
XX
DR   WPI; 2000-273225/24.
DR   N-PSDB; AAA10503.
XX
PT   New mesophilic enzyme from Arthrobacter, used to convert starch
PT   hydrolysate into non-reducing sugars, particularly trehalose, useful as
PT   a sweetener in foods, pharmaceuticals and cosmetics -
XX
PS   Claim 29; Page 47-48; 93pp; English.
XX
CC   This sequence represents a trehalose-releasing enzyme amino acid sequence
CC   from Arthrobacter sp. S34. This enzyme specifically hydrolyses a
CC   non-reducing saccharide having a trehalose structure as an end unit. The
CC   invention relates to this enzyme and also to a non-reducing saccharide
CC   forming enzyme that forms a non-reducing saccharide having a trehalose
CC   structure as an end unit by reducing partial starch hydrolysates. The two
CC   enzymes can be used to produce trehalose, a disaccharide consisting of
CC   two moles of glucose bound at their reducing residues. The disaccharide
CC   is substantially free of reducibility and has moisture retaining
CC   abilities. The enzymes are used to produce non-reducing sugars, they are
CC   particularly used for further enzymatic conversion of starch hydrolysate
CC   to trehalose. The non-reducing sugars produced by the enzymes are useful
CC   in foods, pharmaceuticals  and cosmetics. Non-reducing sugars are
CC   compatible with amino acids and proteins, they do not cause browning, and
CC   have good moisture-retaining properties. The enzymes produce the
CC   non-reducing sugars in high yield and function at a medium temperature
CC   range i.e. between 40 degrees celsius and 60 degrees celsius, and in an
CC   acidic pH range, i.e. a pH of less than 7. Production of non-reducing
CC   sugars using the enzymes eliminates the need to convert a starch
CC   hydrolysate into sugar alcohols (non-reducing) by catalytic
CC   hydrogenation.
XX
SQ   Sequence   575 AA;


  Query Match             31.7%;  Score 936;  DB 21;  Length 575;
  Best Local Similarity   38.7%;  Pred. No. 2.3e-78;
  Matches  198;  Conservative   82;  Mismatches  177;  Indels   54;  Gaps   10;

Qy     15 FKLWAPYVNSVKLKL--SKKLIPMEKNDEGFFEVE--IDDIEENLTYSYIIEDKREIPDP 70
          | :|||    | | :   :  :|: ::: |:: ::   |   : : | |::: |    ||
Db      5 fpvwapqaaqvtlvvgqgraelpltrdengwwalqqpwdggpdlvdygylvdgkgpfadp 64

Qy     71 ASRYQPLGVHDKSQLIRTDYQILDLGKV----------------KIEDLIIYELHVGTFS 114
           |  || |||             :||:                  :   :|||||||||:
Db     65 rslrqprgvh------------elgrefdparyawgddgwrgrdltgaviyelhvgtft 111

Qy    115 QEGNFKGVIEKLDYLKDLGITGIELMPVAQFPGNRDWGYDGVFLYAVQNTYGGPWELAKL 174
           ||     | :||:|  ||:  :||:||  | |   ||||||  |||   ||||    :
Db    112 pegtldsairrldhlvrlgvdavellpvnafngthgwgydgvlwyavhepyggpeayqrf 171

Qy    175 VNEAHKRGIAVILDVVYNHIGPEGNYLLGLGPYFSDRYKTPWGLTFNFDDRGCDQVRKFI 234
          |:  | ||:||: ||||||:|| ||:|   |||         ||   | |    |:||::|
Db    172 vdacharglavvqdvvynhlgpsgnhlpdfgpylgsgaantwgdalnldgplsdevrryi 231

Qy    235 LENVEYWFKTFKIDGLRLDAVHAIFDNSPKHILQEIA----EKAHQLGK--FVIAESDLN 288
          ::|  || :    ||||||||||: |    |:|:|:|    | | :||:   :|||||||
Db    232 idnavywlrdmhadglrldavhalrdaralhlleelaarvdelagelgrpltliaesdln 291

Qy    289 DPKIVKDDC--GYKIDAQWVDDFHHAVHAFITKEKDYYYQDFGRIEDIEKTFKDVFVYDG 346
          |||:::     || :|||| || |||||| :| |   || ||| :  : | |:  : :||
Db    292 dpklirsraahgygldaqwdddvhhavhanvtgetvgyyadfgglgalvkvfqrgwfhdg 351

Qy    347 KYSRYRGRTHGAPVG-DLPPRKFVVFIQNHDQVGNRGNGERLSILTDKTTYLMAATLYIL 405
           :| :| | || |:  |:| |: | | |:|||||||  |:|:|    : :   || | :|
Db    352 twssfrerhhgrpldpdipfrrlvafaqdhdqvgnravgdrmsaqvgegslaaaaalvll 411

Qy    406 SPYIPLIFMGEEYYETNPFFFFSDFSDPVLIKGVREGRLKENNQM-------IDPQSEEA 458
          |: |::|||||:    |: ||:   :| | :    ||: |  :|        |||
Db    412 gpftpmlfmgeewgartpwqfftshpepelgeatargriaefarmgwdpavvpdpqdpat 471

Qy    459 FLKSKLSWKIDEE-----VLDYYKQLINIRK 484
          | :| | |   |      :| :|  || :|:
Db    472 farshldwseperephagllafytdlialrr 502


RESULT 11
AAY85165
ID   AAY85165 standard; Protein; 575 AA.
XX
AC   AAY85165;
XX
DT   23-JUN-2000  (first entry)
XX
DE   Trehalose-releasing enzyme amino acid sequence.
XX
KW   Non-reducing saccharide-forming enzyme; trehalose; disaccharide; food;
KW   high yield; trehalose releasing enzyme; moisture retention; cosmetic;
KW   production.
XX
OS   Arthrobacter sp. S34.
XX
PN   EP990704-A2.
XX
PD   05-APR-2000.
XX
PF   13-SEP-1999;   99EP-0307220.
XX
PR   11-SEP-1998;   98JP-0258394.
PR   11-DEC-1998;   98JP-0352252.
PR   26-JAN-1999;   99JP-0016931.
```

```
XX
PA   (HAYB ) HAYASHIBARA SEIBUTSU KAGAKU.
XX
PI   Yamamoto T,  Maruta K,  Kubota M,  Fukuda S,  Miyake T;
XX
DR   WPI; 2000-273225/24.
DR   N-PSDB; AAA10516.
XX
PT   New mesophilic enzyme from Arthrobacter, used to convert starch
PT   hydrolysate into non-reducing sugars, particularly trehalose, useful as
PT   a sweetener in foods, pharmaceuticals and cosmetics -
XX
PS   Example 6-2; Page 56-58; 93pp; English.
XX
CC   This sequence represents a trehalose-releasing enzyme amino acid sequence
CC   from Arthrobacter sp. S34. This enzyme specifically hydrolyses a
CC   non-reducing saccharide having a trehalose structure as an end unit. The
CC   invention relates to this enzyme and also to a non-reducing saccharide
CC   forming enzyme that forms a non-reducing saccharide having a trehalose
CC   structure as an end unit by reducing partial starch hydrolysates. The two
CC   enzymes can be used to produce trehalose, a disaccharide consisting of
CC   two moles of glucose bound at their reducing residues. The disaccharide
CC   is substantially free of reducibility and has moisture retaining
CC   abilities. The enzymes are used to produce non-reducing sugars, they are
CC   particularly used for further enzymatic conversion of starch hydrolysate
CC   to trehalose. The non-reducing sugars produced by the enzymes are useful
CC   in foods, pharmaceuticals  and cosmetics. Non-reducing sugars are
CC   compatible with amino acids and proteins, they do not cause browning, and
CC   have good moisture-retaining properties. The enzymes produce the
CC   non-reducing sugars in high yield and function at a medium temperature
CC   range i.e. between 40 degrees celsius and 60 degrees celsius, and in an
CC   acidic pH range, i.e. a pH of less than 7. Production of non-reducing
CC   sugars using the enzymes eliminates the need to convert a starch
CC   hydrolysate into sugar alcohols (non-reducing) by catalytic
CC   hydrogenation.
XX
SQ   Sequence   575 AA;


  Query Match             31.7%;  Score 936;  DB 21;  Length 575;
  Best Local Similarity   38.7%;  Pred. No. 2.3e-78;
  Matches  198;  Conservative   82;  Mismatches  177;  Indels   54;  Gaps   10;

Qy      15 FKLWAPYVNSVKLKL--SKKLIPMEKNDEGFFEVE--IDDIEENLTYSYIIEDKREIPDP 70
           | :|||    | | :   :  :|: ::: |:: ::   |   : : | |::: |    ||
Db       5 fpvwapqaaqvtlvvgqgraelpltrdengwwalqqpwdggpdlvdygylvdgkgpfadp 64

Qy      71 ASRYQPLGVHDKSQLIRTDYQILDLGKV----------------KIEDLIIYELHVGTFS 114
             |  || |||             :||:                :   :|||||||||:
Db      65 rslrqprgvh-------------elgrefdparyawgddgwrgrdltgaviyelhvgtft 111

Qy     115 QEGNFKGVIEKLDYLKDLGITGIELMPVAQFPGNRDWGYDGVFLYAVQNTYGGPWELAKL 174
            ||     | :||:|  ||:  :||:||  | |   ||||||  |||   ||||    :
Db     112 pegtldsairrldhlvrlgvdavellpvnafngthgwgydgvlwyavhepyggpeayqrf 171

Qy     175 VNEAHKRGIAVILDVVYNHIGPEGNYLLGLGPYFSDRYKTPWGLTFNFDDRGCDQVRKFI 234
           |:  | ||:||: ||||||:|| ||:|   |||        ||   | |    |:||::|
Db     172 vdacharglavvqdvvynhlgpsgnhlpdfgpylgsgaantwgdalnldgplsdevrryi 231

Qy     235 LENVEYWFKTFKIDGLRLDAVHAIFDNSPKHILQEIA----EKAHQLGK--FVIAESDLN 288
           ::|  || :    ||||||||||: |    |:|:|:|    | | :||:   :|||||||
Db     232 idnavywlrdmhadglrldavhalrdaralhlleelaarvdelagelgrpltliaesdln 291

Qy     289 DPKIVKDDC--GYKIDAQWVDDFHHAVHAFITKEKDYYYQDFGRIEDIEKTFKDVFVYDG 346
           |||:::     || :|||| || |||||| :| |   || ||| :  : | |:  : :||
Db     292 dpklirsraahgygldaqwdddvhhavhanvtgetvgyyadfgglgalvkvfqrgwfhdg 351

Qy     347 KYSRYRGRTHGAPVG-DLPPRKFVVFIQNHDQVGNRGNGERLSILTDKTTYLMAATLYIL 405
            :| :| | || |:  |:| |: | | |:|||||||  |:|:|    : :   || | :|
Db     352 twssfrerhhgrpldpdipfrrlvafaqdhdqvgnravgdrmsaqvgegslaaaaalvll 411

Qy     406 SPYIPLIFMGEEYYETNPFFFFSDFSDPVLIKGVREGRLKENNQM-------IDPQSEEA 458
            |: |::|||||:    |: ||:   :| | :    ||: |  :|        |||
Db     412 gpftpmlfmgeewgartpwqfftshpepelgeatargriaefarmgwdpavvpdpqdpat 471

Qy     459 FLKSKLSWKIDEE-----VLDYYKQLINIRK 484
           | :| | |   |      :| :|  || :|:
Db     472 farshldwseperephagllafytdlialrr 502


RESULT  12
AAR80037
ID   AAR80037 standard; Protein; 652 AA.
XX
AC   AAR80037;
XX
DT   11-JUL-1996  (first entry)
XX
DE   Bacillus stearothermophilus thermostable branching enzyme.
XX
KW   Thermostable; 1,4-alpha-glucan; branching enzyme; amylose;
KW   amylopectin; transferase.
XX
OS   Bacillus stearothermophilus (strain TRBE14).
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1
FT                   /note= "Met residue corresponds to TTG codon"
XX
PN   JP07213287-A.
XX
PD   15-AUG-1995.
XX
PF   04-NOV-1993;   93JP-0311004.
XX
PR   04-NOV-1993;   93JP-0311004.
XX
PA   (EZAK ) EZAKI GLICO CO.
XX
DR   WPI; 1995-315926/41.
DR   N-PSDB; AAT04542.
XX
PT   Thermostable branching enzyme gene and recombinant plasmid contg. it
PT    - used to produced a thermostable branching enzyme
XX
PS   Claim 3; Page 6-12; 14pp; Japanese.
XX
CC   A gene coding for a thermostable branching enzyme has been isolated
CC   from chromosomal DNA of B.stearothermophilus strain TRBE14 and
CC   sequenced (see AAT04542). The enzyme (AAR80037) was found to have an
CC   activity of 80% or more at up to 65 deg.C
XX
SQ   Sequence   652 AA;


  Query Match             12.2%;  Score 360;  DB 16;  Length 652;
  Best Local Similarity   25.3%;  Pred. No. 1.7e-24;
  Matches  142;  Conservative   72;  Mismatches  186;  Indels  162;  Gaps   25;

Qy       4 FGGNIEKNKG----IFKLWAPYVNSVKL-----------KLSKKLIPMEKNDEGFFEVE 47
           || ::    |     | :|||:   |:|           :| |       |||| : :
Db      26 fgahvineggkvgtrfcvwapharevrlvgsfndwdgtdfrlek------vndegvwtiv 79

Qy      48 IDDIEENLTYSY-IIEDKREI---PDPASRYQPLGVHDKSQLIRTD---YQILDLG---- 96
           : :  |   | | |:    ::    || : |  |  |  |  | |   ||  |
Db      80 vpenleghlykyeivtpdgqvlfkadpyafyselrphtas--iaydlkgyqwndqswkrk 137

Qy      97 ----KVKIEDLIIYELHVGTF-SQEGNFKGVIEKLD----YLKDLGITGIELMPVAQFPG 147
               ::  : ::|||||  |::  ::| |    |  |    |: | | | |||:|: : |
Db     138 krrkriydqpmviyelhfgswkkkdgrfytyremadelisyvldhgfthiellplvehpl 197

Qy     148 NRDWGYDGVFLYAVQNTYGGPWELAKLVNEAHKRGIAVILDVVYNHI--GPEGNYLLGLG 205
           :| ||| |   ||| : || | :    |:  |: || ||:| |  |      | |:
Db     198 drswgyqgtgyyavtsrygtphdfmyfvdrchqagigvimdwvpghfckdahglymfdga 257

Qy     206 PYF-----SDRYKTPWGLTFNFDDRGCDQVRKFILENVEYWFKTFKIDGLRLDAV----- 255
```

```
          | :      ||    || | || | |  :|| |:: |  :| : : ||| |:|||
Db    258 ptyeyanekdrenyvwg-tanf-dlgkpevrsflisnalfwleyyhidgfrvdavanmly 315

Qy    256 ----HAIFDNS-PKHILQEIAEK--AHQLGKFVIAESDLNDPKIVKD------DCGYKID 302
              :::|      |::: |   |:     :|||   : ||:            || :
Db    316 wpnndrlyenpyaveflrklneavfaydpnalmiaedstdwpkvtaptyegglgfnykwn 375

Qy    303 AQWVDDF----------HHAVHAFITKEKDYYYQDFGRIEDIEKTFKDVFVYDGKYSRYR 352
            |::|             ||  :|    | |      |:    |    |  || |
Db    376 mgwmndmlkymetppyerrhvhnqvtfslyays-----enfilpfshdevvhgkks--- 427

Qy    353 GRTHGAPVGDLPPRKFVVFIQNHDQVGNRGNG---ERLSILTDKTTYLMAATLYILSPYI 409
                                  : |:  |   |: : |     |:||       |
Db    428 ------------------------llnkmpgsyeekfaqlrllygymma------hpgk 456

Qy    410 PLIFMGEEYYETNPFFFFSDFSDPVLIKGVREGRLKENNQMIDPQSEEAFLKSKLSW--- 466
           |:||| |:                                    | :|   : :| |
Db    457 kllfmgnef---------------------------------aqfdewkfedeldwvlf 482

Qy    467 --KIDEEVLDYYKQLINIRKRY 486
            ::  :: || |:||   |||
Db    483 dfelhrkmndymkeliacykry 504
```

RESULT 13
AAR96109
ID AAR96109 standard; Protein; 652 AA.
XX
AC AAR96109;
XX
DT 20-AUG-1996 (first entry)
XX
DE Starch-branching-enzyme.
XX
KW Starch-branching-enzyme; thermostable; Q-enzyme; Escherichia coli;
KW starch; amylopectin; glucan; cyclodextrin; food; food-additive;
KW adhesive; biodegradable plastic.
XX
OS Bacillus stearothermophilus strain TRBE14 (FERM P-13916).
XX
FH Key Location/Qualifiers
FT Peptide 238..249
FT /note= "Conserved amylase sequence used to
FT construct primer AAT27126"
FT Peptide 309..317
FT /note= "Conserved amylase sequence"
FT Peptide 414..423
FT /note= "Conserved amylase sequence used to
FT construct primer AAT27127"
XX
PN EP710674-A2.
XX
PD 08-MAY-1996.
XX
PF 13-SEP-1995; 95EP-0250222.
XX
PR 31-JUL-1995; 95JP-0195647.
PR 13-SEP-1994; 94JP-0218554.
XX
PA (EZAK) EZAKI GLICO CO LTD.
XX
PI Fujii K, Imanaka T, Nakamura H, Okada S, Takaha T;
PI Takata H, Terada Y, Yanase M;
XX
DR WPI; 1996-223407/23.
DR N-PSDB; AAT27128.
XX
PT New glucan with cyclic, branched core and outer branched region -
PT useful in foods, drinks, adhesives, infusion compsns. etc., readily
PT soluble in water forming solns. of low viscosity
XX
PS Example A-2; Page 28-31; 50pp; English.
XX
CC This sequence represents a thermostable starch- branching-enzyme
CC (Q-enzyme, EC-2.4.1.18) from Bacillus stearothermophilus TRBE14
CC (FERM P-13916). Conserved regions present in amylases are used to
CC construct primers AAT27126-7, which are used in PCR to amplify a
CC fragment of the gene, which is then used as a probe to isolate a
CC full-length clone in Escherichia coli. The recombinant enzyme may
CC be reacted with starch or amylopectin to produce a glucan with a
CC degree of polymerisation of at least 50, with an inner branched
CC cyclic region and an outer branched region. The product is freely
CC soluble in water, forming a low-viscosity solution, is more
CC resistance to degradation than native starch, does not react with
CC proteins and amino acids, and is degraded to glucose in vivo, so is
CC easily digestible. The glucan may be used in foods, beverages,
CC food-additives, adhesives, in cyclodextrin production and as a
CC starch substitute in biodegradable plastic.
XX
SQ Sequence 652 AA;

Query Match 12.2%; Score 360; DB 17; Length 652;
Best Local Similarity 25.3%; Pred. No. 1.7e-24;
Matches 142; Conservative 72; Mismatches 186; Indels 162; Gaps 25;

```
Qy      4 FGGNIEKNKG----IFKLWAPYVNSVKL-----------KLSKKLIPMEKNDEGFFEVE 47
          || ::    |     | :|||:   |:|           :| |       |||| : :
Db     26 fgahvineggkvgtrfcvwapharevrlvgsfndwdgtdfrlek------vndegvwtiv 79

Qy     48 IDDIEENLTYSY-IIEDKREI---PDPASRYQPLGVHDKSQLIRTD---YQILDLG---- 96
          : :  |   | | |:    ::    || : |  |  |  |  |  |   || |
Db     80 vpenleghlykyeivtpdgqvlfkadpyafyselrphtas--iaydlkgyqwndqswkrk 137

Qy     97 ----KVKIEDLIIYELHVGTF-SQEGNFKGVIEKLD----YLKDLGITGIELMPVAQFPG 147
               ::  : ::|||| |::  ::| |    |  |    |: | | | |||:|: : |
Db    138 krrkriydqpmviyelhfgswkkkdgrfytyremadelisyvldhgfthiellplvehpl 197

Qy    148 NRDWGYDGVFLYAVQNTYGGPWELAKLVNEAHKRGIAVILDVVYNHI--GPEGNYLLGLG 205
          :| ||| |   ||| : || | :    |:  |: || ||:| |  |      | |:
Db    198 drswgyqgtgyyavtsrygtphdfmyfvdrchqagigvimdwvpghfckdahglymfdga 257

Qy    206 PYF-----SDRYKTPWGLTFNFDDRGCDQVRKFILENVEYWFKTFKIDGLRLDAV----- 255
          | :      ||    || | || | |  :|| |:: |  :| : : ||| |:|||
Db    258 ptyeyanekdrenyvwg-tanf-dlgkpevrsflisnalfwleyyhidgfrvdavanmly 315

Qy    256 ----HAIFDNS-PKHILQEIAEK--AHQLGKFVIAESDLNDPKIVKD------DCGYKID 302
               :::|      |::: |   |:     :|||   : ||:            || :
Db    316 wpnndrlyenpyaveflrklneavfaydpnalmiaedstdwpkvtaptyegglgfnykwn 375

Qy    303 AQWVDDF----------HHAVHAFITKEKDYYYQDFGRIEDIEKTFKDVFVYDGKYSRYR 352
            |::|             ||  :|    | |      |:    |    |  || |
Db    376 mgwmndmlkymetppyerrhvhnqvtfslyays-----enfilpfshdevvhgkks--- 427

Qy    353 GRTHGAPVGDLPPRKFVVFIQNHDQVGNRGNG---ERLSILTDKTTYLMAATLYILSPYI 409
                                  : |:  |   |: : |     |:||       |
Db    428 ------------------------llnkmpgsyeekfaqlrllygymma------hpgk 456

Qy    410 PLIFMGEEYYETNPFFFFSDFSDPVLIKGVREGRLKENNQMIDPQSEEAFLKSKLSW--- 466
           |:||| |:                                    | :|   : :| |
Db    457 kllfmgnef---------------------------------aqfdewkfedeldwvlf 482

Qy    467 --KIDEEVLDYYKQLINIRKRY 486
            ::  :: || |:||   |||
Db    483 dfelhrkmndymkeliacykry 504
```

RESULT 14
AAR23787
ID AAR23787 standard; Protein; 658 AA.
XX
AC AAR23787;
XX
DT 04-NOV-1992 (first entry)
XX
DE Heat-resistant pullanase.

```
XX
KW   Enzyme; pTPI.
XX
OS   Synthetic.
XX
PN   JP04099489-A.
XX
PD   31-MAR-1992.
XX
PF   27-MAR-1990;   90JP-0077230.
XX
PR   27-MAR-1990;   90JP-0077230.
XX
PA   (EZAK ) EZAKI GLICO CO.
PA   (IATR ) IATRON LAB INC.
XX
DR   WPI; 1992-157263/19.
DR   P-PSDB; AAR23787.
XX
PT   Heat-resistant pullulanase gene - present recombinant plasmid are new
PT   and produced in large amts. using a proper host
XX
PS   Claim 3; Page 2-3; 8pp; Japanese.
XX
CC   The sequence encoding this protein may be contained in a recombinant
CC   plasmid.  The heat-resistant pullanase can be produced in large amts.
CC   using a suitable host.
CC   Plasmid pTPI contg. the pullanase gene (Japan Patent Apppln. No.
CC   174037/88) was cleaved by various restriction enzymes and inserted
CC   into plasmid pV118/119.  Single-stranded DNA was prepd. using helper
CC   phage M13K07 and the base sequence determined by the dideoxy method.
XX
SQ   Sequence   658 AA;
```

```
  Query Match             12.1%;  Score 356;  DB 13;  Length 658;
  Best Local Similarity   22.9%;  Pred. No. 4.1e-24;
  Matches  141;  Conservative   99;  Mismatches  200;  Indels  176;  Gaps   26;

Qy       2 FSFGGNI----EKNKGIFKLWAPYVNSVKLKL------SKKLIPMEKNDEGFFEVEIDDI 51
           | : |::     : | |||:|||    ||:||        : | :|: ::| :   :
Db      44 ffydgdlgvaytrektifkvwaptatevkvklidpktekaeyialerlekgvwtttvfgd 103

Qy      52 EENLTYSYII---EDKREIPDPASRYQPLGVHDKSQLIRTDY-QILDLGKVKIE------ 101
            |   |:|::      ||  ||         :  :  :  :|  |:|| | : |
Db     104 megryytyvvcvnavwreavdp---------yavavsvngeygviidlaktrtekppipp 154

Qy     102 -----DLIIYELHVGTF--------SQEGNFKGVIEK-----------LDYLKDLGITGI 137
                | :|||:|:  |        : :| : |: |           | |||:||:| :
Db     155 itspadaviyemhirdftihpesgvthkgkylgltepetkgpnntptglsylkelgvthv 214

Qy     138 ELMPVAQFPG--------NRDWGYDGVFLYAVQNTY----GGPW----ELAKLVNEAHKR 181
           ||:|   | |          :|||: :   | : :|      |:    || : ::
Db     215 ellpfndfagvdernplqqynwgynplhynapegsyatdpndpyarirelkraihtlqsn 274

Qy     182 GIAVILDVVYNHIGPEGNYL--------LGLGPYFS-DRYKTPWGLTFNFDDRGCDQ--V 230
           || ||:||||||:     |:        :  | ||  | |  |   |   :|   ::  |
Db     275 girvimdvvynhv-----yirdqssfekivpgyyfrydaygnpsngtgvgndiaserkmv 329

Qy     231 RKFILENVEYWFKTFKIDGLRLDAVHAIFDNSPKHILQEIAEKAHQLGKFVIAESDLNDP 290
           ||:|:::| :| : : ::| | | :  | |      ::|:                |  ||
Db     330 rkwiidsvrfwveeyhvngfrfdlm-gildvetmkavreml-------------dtldp 374

Qy     291 KIVKDDCGYKIDAQWVDDFHHAVHAFITKEKDYYYQDFGRIEDIEKTFKDVFVYDGKYSR 350
            |:    |: :               :: |:    |:  ::  |   | | |    | |
Db     375 silvfgegwdlptp-----------lsseqkatmqnaeklpri-gyfndyfrdsvkgst 421

Qy     351 YRGRTHGAPVGD---------------------LPPRKFVVFIQNHD--------QVGN 380
           ::    |  :||                     | | : : ::::||        :: |
Db     422 fqlpdrgfalgdsddrekvkvaisgsigkknglflhptqtinyveshdnhtfwdkmeian 481

Qy     381 RGNGERLSILTDKTTYLMAATLYILSPYIPLIFMGEEYYETNPFFFFSDFSDPVLIKGVR 440
               |:    | |    :|  : :||  || :  |:|:| |               :|:
Db     482 gheseq----trkkrqklataivllsqgipflhsgqefyrtk-------------qgi- 522

Qy     441 EGRLKENNQMIDPQSEEAFLKSKLSWKIDEEVLD---YYKQLINIRKRYNNCKRVKEVRR 497
                 :|    | |     ::: ||   |  |   | : || ||| :   :   |
Db     523 ------dnsynapds-----inqidwkrkseheddvryvqglirirkshrafrftteaei 571

Qy     498 EGNCITLIMEKIGIIA 513
             :   |      |||
Db     572 nnhlffleptpasiia 587
```

```
RESULT  15
AAR11271
ID   AAR11271 standard; Protein; 639 AA.
XX
AC   AAR11271;
XX
DT   29-MAY-1991  (first entry)
XX
DE   B.stearothermophilus branching enzyme.
XX
KW   Branching enzyme; amylaceous substances; transferase;
KW   thermostability.
XX
OS   Bacillus stearothermophilus.
XX
PN   EP418945-A.
XX
PD   27-MAR-1991.
XX
PF   22-AUG-1990;   90EP-0202264.
XX
PR   23-AUG-1989;   89NL-0002128.
XX
PA   (COOP-) COOP VERKOOP AVEBE.
XX
PI   Kiel JAKW,  Boels JM,  Beldman G,  Venema G;
XX
DR   WPI; 1991-088587/13.
DR   N-PSDB; AAQ11072.
XX
PT   Branching enzyme obtd. from Bacillus stearothermophilus - has good
PT   thermostability, for use in modifying starch-like substances
XX
PS   Disclosure; fig 4; 13pp; English.
XX
CC   This branching enzyme is derived from B.stearothermophilus and is
CC   encoded by the glgB gene.  It is stable at high temps. and can be
CC   produced in large amts. using recombinant techniques.  It is a
CC   transferase allowing formation of alpha(1-6) branches in starch-like
CC   substances.  It is useful for extending the shelf life of such
CC   materials and can be used in the prepn. of human and animal foods.
XX
SQ   Sequence   639 AA;
```

```
  Query Match             11.6%;  Score 342.5;  DB 12;  Length 639;
  Best Local Similarity   24.9%;  Pred. No. 7.2e-23;
  Matches  135;  Conservative   89;  Mismatches  197;  Indels  121;  Gaps   26;

Qy       4 FGGNIEKNKGI----FKLWAPYVNSVKLKLS------KKLIPMEKNDEGFFEVEIDDIEE 53
           || :: |  |:    | :|||:   |:|  |           |: :::| : : | :  |
Db      26 fgahvikkngmvgtrfcvwapharevrlvgsfnewngtnfnlmkvsnqgvwmifipenle 85

Qy      54 NLTYSYII--EDKREI--PDPASRYQPLGVHDKSQLIR-TDYQILD--------LGKVKI 100
             | | |   |   :   || : |  |  |  | :     ||  |          ::
Db      86 ghlykyeittndgnvllksdpyafyselrphtasivynikgyqwndqtwrrkkqrkriyd 145

Qy     101 EDLIIYELHVGTF--SQEGNF----KGVIEKLDYLKDLGITGIELMPVAQFPGNRDWGYD 154
           : | ||||| |::   ::|:|    :   | : |: : | | |||:|: : | :| |||
Db     146 qplfiyelhfgswkkkedgsfytyqemaeelipyvlehgfthiellplvehpfdrswgyq 205

Qy     155 GVFLYAVQNTYGGPWELAKLVNEAHKRGIAVILDVVYNHI--GPEGNYLLGLGPYF---- 208
```

```
          |:  |:  : || | :|   ::  |: || |||| |  |      | |:    | :
Db    206 gigyysatsrygtphdlmyfidrchqagigvildwvpghfckdshglymfdgapayeyan 265

Qy    209 -SDRYKTPWGLTFNFDDRGCDQVRKFILENVEYWFKTFKIDGLRLDAV---------HAI 258
            ||     || | || | |  :|| |:: |  :| : | :|| |:|||          :
Db    266 mqdrenyvwg-tanf-dlgkpevrsflisnalfwmeyfhvdgfrvdavanmlywpnsdvl 323

Qy    259 FDNS-PKHILQEIAEK--AHQLGKFVIAESDLNDPKIVKD--DCG----YKIDAQWVDDF 309
          : |:     ||:: |   |:     :|||   : |::     | |    || :  |::|
Db    324 ykntyaveflqklnetvfaydpnilmiaedstdwprvtaptydgglgfnykwnmgwmndi 383

Qy    310 HHAVHAFITKEKDYYYQDFGRIEDIEKTFKDVFVYDGKYSRYRGRTHGAPVGDLPPRKFV 369
              :     |   |  :       : ||  :: |   :
Db    384 --ltymetppehrkyvhn-------kvtfsllyaysenf--------------------- 413

Qy    370 VFIQNHDQVGNRGNGERLSIL----TDKTTYLMAATLYILS-PYIPLIFMGEEYYETNPF 424
          :   :||:| : |    || :     :|   |     |:|: |   |:||| |: :
Db    414 ilpfshdevvh-gkksllskmpgtyeekfaqlrllygyllthpgkkllfmggefgq---- 468

Qy    425 FFFSDFSDPVLIKGVREGRLKENNQMIDPQSEEAFLKSKLSWKIDEEVLDYYKQLINIRK 484
            | :: |           |:: : |:              : :   :  | |:|:   |
Db    469 --fdewkd----------leqldwml-----------fdfdmhrnmnmyvkellkcyk 503

Qy    485 RY 486
          ||
Db    504 ry 505
```

Search completed: July 15, 2002, 12:17:34
Job time: 71 sec

GenCore version 4.5

OM protein - protein search, using sw model

Run on: July 15, 2002, 12:18:04 ; Search time 25.09 Seconds
(without alignments)
541.277 Million cell updates/sec

Title: US-09-298-924-8
Perfect score: 2954
Sequence: 1 MFSFGGNIEKNKGIFKLWAP.....KLKKDELIKVNRGVGVYQLE 556

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2954	100.0	556	2	US-08-505-377-1	Sequence 1, Appli
2	2954	100.0	556	3	US-08-798-269-1	Sequence 1, Appli
3	2954	100.0	556	4	US-09-055-210-1	Sequence 1, Appli
4	1748.5	59.2	559	4	US-09-242-690A-15	Sequence 15, Appl
5	969	32.8	589	1	US-08-399-646-2	Sequence 2, Appli
6	969	32.8	589	1	US-08-607-321-2	Sequence 2, Appli
7	969	32.8	589	2	US-08-961-240-2	Sequence 2, Appli
8	969	32.8	589	2	US-08-605-501-2	Sequence 2, Appli
9	969	32.8	596	1	US-08-399-646-12	Sequence 12, Appl
10	969	32.8	596	1	US-08-607-321-12	Sequence 12, Appl
11	969	32.8	596	2	US-08-961-240-12	Sequence 12, Appl
12	969	32.8	596	2	US-08-605-501-12	Sequence 12, Appl
13	959	32.5	597	1	US-08-399-646-4	Sequence 4, Appli
14	959	32.5	597	1	US-08-607-321-4	Sequence 4, Appli
15	959	32.5	597	2	US-08-961-240-4	Sequence 4, Appli
16	959	32.5	597	2	US-08-605-501-4	Sequence 4, Appli
17	959	32.5	598	1	US-08-399-646-14	Sequence 14, Appl
18	959	32.5	598	1	US-08-607-321-14	Sequence 14, Appl
19	959	32.5	598	2	US-08-961-240-14	Sequence 14, Appl
20	959	32.5	598	2	US-08-605-501-14	Sequence 14, Appl
21	360	12.2	652	4	US-08-528-026C-4	Sequence 4, Appli
22	319.5	10.8	718	4	US-09-346-237-10	Sequence 10, Appl
23	310.5	10.5	606	4	US-09-187-124-2	Sequence 2, Appli
24	300.5	10.2	740	2	US-08-410-784A-4	Sequence 4, Appli
25	298	10.1	713	4	US-09-346-237-9	Sequence 9, Appli
26	296.5	10.0	818	2	US-08-410-784A-2	Sequence 2, Appli
27	296.5	10.0	818	4	US-09-346-237-11	Sequence 11, Appl
28	292	9.9	829	4	US-09-514-599-6	Sequence 6, Appli
29	292	9.9	862	4	US-09-346-237-1	Sequence 1, Appli
30	284	9.6	893	4	US-09-514-302-4	Sequence 4, Appli
31	284	9.6	1938	4	US-09-514-302-2	Sequence 2, Appli
32	276	9.3	921	4	US-09-514-599-2	Sequence 2, Appli
33	268	9.1	915	4	US-09-346-237-2	Sequence 2, Appli
34	268	9.1	928	4	US-09-514-599-4	Sequence 4, Appli
35	264.5	9.0	820	4	US-09-313-677-21	Sequence 21, Appl
36	264.5	9.0	926	4	US-09-313-677-2	Sequence 2, Appli
37	264.5	9.0	933	4	US-09-313-677-19	Sequence 19, Appl
38	264.5	9.0	967	4	US-09-313-677-17	Sequence 17, Appl
39	250	8.5	928	1	US-08-474-140-11	Sequence 11, Appl
40	250	8.5	928	1	US-08-477-630-11	Sequence 11, Appl
41	250	8.5	928	1	US-08-472-293-11	Sequence 11, Appl
42	250	8.5	928	1	US-08-474-545-11	Sequence 11, Appl
43	250	8.5	928	2	US-08-478-341-11	Sequence 11, Appl
44	250	8.5	928	3	US-08-996-733-11	Sequence 11, Appl
45	249	8.4	588	2	US-08-339-715A-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-08-505-377-1
; Sequence 1, Application US/08505377
; Patent No. 5856146
; GENERAL INFORMATION:
; APPLICANT: MITSUZUMI, Hitoshi
; APPLICANT: KUBOTA, Michio
; APPLICANT: SUGIMOTO, Toshiyuki
; TITLE OF INVENTION: RECOMBINANT THERMOSTABLE ENZYME WHICH
; TITLE OF INVENTION: RELEASES TREHALOSE FROM NON-REDUCING SACCHARIDE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Browdy and Neimark
; STREET: 419 Seventh Street N.W. Ste. 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/505,377
; FILING DATE: 21-JUL-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 190180/1994
; FILING DATE: 21-JUL-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 109128/1995
; FILING DATE: 11-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP NOT YET RECEIVED
; FILING DATE: 04-JUL-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Browdy, Roger L
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: MITSUZUMI=1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-5197
; TELEFAX: (202) 737-3528
; TELEX: 249688
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 556 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein

claims polypep.

US-08-505-377-1

Query Match 100.0%; Score 2954; DB 2; Length 556;
Best Local Similarity 100.0%; Pred. No. 1.1e-280;
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 MFSFGGNIEKNKGIFKLWAPYVNSVKLKLSKKLIPMEKNDEGFFEVEIDDIEENLTYSYI 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MFSFGGNIEKNKGIFKLWAPYVNSVKLKLSKKLIPMEKNDEGFFEVEIDDIEENLTYSYI 60

Qy      61 IEDKREIPDPASRYQPLGVHDKSQLIRTDYQILDLGKVKIEDLIIYELHVGTFSQEGNFK 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 IEDKREIPDPASRYQPLGVHDKSQLIRTDYQILDLGKVKIEDLIIYELHVGTFSQEGNFK 120

Qy     121 GVIEKLDYLKDLGITGIELMPVAQFPGNRDWGYDGVFLYAVQNTYGGPWELAKLVNEAHK 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 GVIEKLDYLKDLGITGIELMPVAQFPGNRDWGYDGVFLYAVQNTYGGPWELAKLVNEAHK 180

Qy     181 RGIAVILDVVYNHIGPEGNYLLGLGPYFSDRYKTPWGLTFNFDDRGCDQVRKFILENVEY 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 RGIAVILDVVYNHIGPEGNYLLGLGPYFSDRYKTPWGLTFNFDDRGCDQVRKFILENVEY 240

Qy     241 WFKTFKIDGLRLDAVHAIFDNSPKHILQEIAEKAHQLGKFVIAESDLNDPKIVKDDCGYK 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 WFKTFKIDGLRLDAVHAIFDNSPKHILQEIAEKAHQLGKFVIAESDLNDPKIVKDDCGYK 300

Qy     301 IDAQWVDDFHHAVHAFITKEKDYYYQDFGRIEDIEKTFKDVFVYDGKYSRYRGRTHGAPV 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 IDAQWVDDFHHAVHAFITKEKDYYYQDFGRIEDIEKTFKDVFVYDGKYSRYRGRTHGAPV 360

Qy     361 GDLPPRKFVVFIQNHDQVGNRGNGERLSILTDKTTYLMAATLYILSPYIPLIFMGEEYYE 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 GDLPPRKFVVFIQNHDQVGNRGNGERLSILTDKTTYLMAATLYILSPYIPLIFMGEEYYE 420

Qy     421 TNPFFFFSDFSDPVLIKGVREGRLKENNQMIDPQSEEAFLKSKLSWKIDEEVLDYYKQLI 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 TNPFFFFSDFSDPVLIKGVREGRLKENNQMIDPQSEEAFLKSKLSWKIDEEVLDYYKQLI 480

Qy     481 NIRKRYNNCKRVKEVRREGNCITLIMEKIGIIASFDDIVINSKITGNLLIGIGFPKKLKK 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 NIRKRYNNCKRVKEVRREGNCITLIMEKIGIIASFDDIVINSKITGNLLIGIGFPKKLKK 540

Qy     541 DELIKVNRGVGVYQLE 556
           ||||||||||||||||
Db     541 DELIKVNRGVGVYQLE 556
```

RESULT 2
US-08-798-269-1
; Sequence 1, Application US/08798269
; Patent No. 6027918
; GENERAL INFORMATION:
; APPLICANT: MITSUZUMI, Hitoshi
; APPLICANT: KUBOTA, Michio
; APPLICANT: SUGIMOTO, Toshiyuki
; TITLE OF INVENTION: RECOMBINANT THERMOSTABLE ENZYME WHICH
; TITLE OF INVENTION: RELEASES TREHALOSE FROM NON-REDUCING SACCHARIDE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Browdy and Neimark
; STREET: 419 Seventh Street N.W. Ste. 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/798,269
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/505,377
; FILING DATE: 21-JUL-1995
; APPLICATION NUMBER: JP 190180/1994
; FILING DATE: 21-JUL-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 109128/1995
; FILING DATE: 11-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP NOT YET RECEIVED
; FILING DATE: 04-JUL-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Browdy, Roger L
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: MITSUZUMI=1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-5197
; TELEFAX: (202) 737-3528
; TELEX: 249688
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 556 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-798-269-1

Query Match 100.0%; Score 2954; DB 3; Length 556;
Best Local Similarity 100.0%; Pred. No. 1.1e-280;
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 MFSFGGNIEKNKGIFKLWAPYVNSVKLKLSKKLIPMEKNDEGFFEVEIDDIEENLTYSYI 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MFSFGGNIEKNKGIFKLWAPYVNSVKLKLSKKLIPMEKNDEGFFEVEIDDIEENLTYSYI 60

Qy      61 IEDKREIPDPASRYQPLGVHDKSQLIRTDYQILDLGKVKIEDLIIYELHVGTFSQEGNFK 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 IEDKREIPDPASRYQPLGVHDKSQLIRTDYQILDLGKVKIEDLIIYELHVGTFSQEGNFK 120

Qy     121 GVIEKLDYLKDLGITGIELMPVAQFPGNRDWGYDGVFLYAVQNTYGGPWELAKLVNEAHK 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 GVIEKLDYLKDLGITGIELMPVAQFPGNRDWGYDGVFLYAVQNTYGGPWELAKLVNEAHK 180

Qy     181 RGIAVILDVVYNHIGPEGNYLLGLGPYFSDRYKTPWGLTFNFDDRGCDQVRKFILENVEY 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 RGIAVILDVVYNHIGPEGNYLLGLGPYFSDRYKTPWGLTFNFDDRGCDQVRKFILENVEY 240

Qy     241 WFKTFKIDGLRLDAVHAIFDNSPKHILQEIAEKAHQLGKFVIAESDLNDPKIVKDDCGYK 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 WFKTFKIDGLRLDAVHAIFDNSPKHILQEIAEKAHQLGKFVIAESDLNDPKIVKDDCGYK 300

Qy     301 IDAQWVDDFHHAVHAFITKEKDYYYQDFGRIEDIEKTFKDVFVYDGKYSRYRGRTHGAPV 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 IDAQWVDDFHHAVHAFITKEKDYYYQDFGRIEDIEKTFKDVFVYDGKYSRYRGRTHGAPV 360

Qy     361 GDLPPRKFVVFIQNHDQVGNRGNGERLSILTDKTTYLMAATLYILSPYIPLIFMGEEYYE 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 GDLPPRKFVVFIQNHDQVGNRGNGERLSILTDKTTYLMAATLYILSPYIPLIFMGEEYYE 420

Qy     421 TNPFFFFSDFSDPVLIKGVREGRLKENNQMIDPQSEEAFLKSKLSWKIDEEVLDYYKQLI 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 TNPFFFFSDFSDPVLIKGVREGRLKENNQMIDPQSEEAFLKSKLSWKIDEEVLDYYKQLI 480

Qy     481 NIRKRYNNCKRVKEVRREGNCITLIMEKIGIIASFDDIVINSKITGNLLIGIGFPKKLKK 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 NIRKRYNNCKRVKEVRREGNCITLIMEKIGIIASFDDIVINSKITGNLLIGIGFPKKLKK 540

Qy     541 DELIKVNRGVGVYQLE 556
           ||||||||||||||||
Db     541 DELIKVNRGVGVYQLE 556
```

```
RESULT   3
US-09-055-210-1
; Sequence 1, Application US/09055210
; Patent No. 6346394
;  GENERAL INFORMATION:
;    APPLICANT:  MITSUZUMI, Hitoshi
;    APPLICANT:  KUBOTA, Michio
;    APPLICANT:  SUGIMOTO, Toshiyuki
;    TITLE OF INVENTION:  RECOMBINANT THERMOSTABLE ENZYME WHICH
;    TITLE OF INVENTION:  RELEASES TREHALOSE FROM NON-REDUCING SACCHARIDE
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Browdy and Neimark
;      STREET:  419 Seventh Street N.W. Ste. 300
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/055,210
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/505,377
;      FILING DATE:  21-JUL-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 109128/1995
;      FILING DATE:  11-APR-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP NOT YET RECEIVED
;      FILING DATE:  04-JUL-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Browdy, Roger L
;      REGISTRATION NUMBER:  25,618
;      REFERENCE/DOCKET NUMBER:  MITSUZUMI-1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 628-5197
;      TELEFAX:  (202) 737-3528
;      TELEX:  249688
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  556 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-055-210-1
```

Cleaved drawing to DNA w/ T.D.

```
  Query Match           100.0%;  Score 2954;  DB 4;  Length 556;
  Best Local Similarity  100.0%;  Pred. No. 1.1e-280;
  Matches  556;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MFSFGGNIEKNKGIFKLWAPYVNSVKLKLSKKLIPMEKNDEGFFEVEIDDIEENLTYSYI 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MFSFGGNIEKNKGIFKLWAPYVNSVKLKLSKKLIPMEKNDEGFFEVEIDDIEENLTYSYI 60

Qy      61 IEDKREIPDPASRYQPLGVHDKSQLIRTDYQILDLGKVKIEDLIIYELHVGTFSQEGNFK 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 IEDKREIPDPASRYQPLGVHDKSQLIRTDYQILDLGKVKIEDLIIYELHVGTFSQEGNFK 120

Qy     121 GVIEKLDYLKDLGITGIELMPVAQFPGNRDWGYDGVFLYAVQNTYGGPWELAKLVNEAHK 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 GVIEKLDYLKDLGITGIELMPVAQFPGNRDWGYDGVFLYAVQNTYGGPWELAKLVNEAHK 180

Qy     181 RGIAVILDVVYNHIGPEGNYLLGLGPYFSDRYKTPWGLTFNFDDRGCDQVRKFILENVEY 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 RGIAVILDVVYNHIGPEGNYLLGLGPYFSDRYKTPWGLTFNFDDRGCDQVRKFILENVEY 240

Qy     241 WFKTFKIDGLRLDAVHAIFDNSPKHILQEIAEKAHQLGKFVIAESDLNDPKIVKDDCGYK 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 WFKTFKIDGLRLDAVHAIFDNSPKHILQEIAEKAHQLGKFVIAESDLNDPKIVKDDCGYK 300

Qy     301 IDAQWVDDFHHAVHAFITKEKDYYYQDFGRIEDIEKTFKDVFVYDGKYSRYRGRTHGAPV 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 IDAQWVDDFHHAVHAFITKEKDYYYQDFGRIEDIEKTFKDVFVYDGKYSRYRGRTHGAPV 360

Qy     361 GDLPPRKFVVFIQNHDQVGNRGNGERLSILTDKTTYLMAATLYILSPYIPLIFMGEEYYE 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 GDLPPRKFVVFIQNHDQVGNRGNGERLSILTDKTTYLMAATLYILSPYIPLIFMGEEYYE 420

Qy     421 TNPFFFFSDFSDPVLIKGVREGRLKENNQMIDPQSEEAFLKSKLSWKIDEEVLDYYKQLI 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 TNPFFFFSDFSDPVLIKGVREGRLKENNQMIDPQSEEAFLKSKLSWKIDEEVLDYYKQLI 480

Qy     481 NIRKRYNNCKRVKEVRREGNCITLIMEKIGIIASFDDIVINSKITGNLLIGIGFPKKLKK 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 NIRKRYNNCKRVKEVRREGNCITLIMEKIGIIASFDDIVINSKITGNLLIGIGFPKKLKK 540

Qy     541 DELIKVNRGVGVYQLE 556
           ||||||||||||||||
Db     541 DELIKVNRGVGVYQLE 556
```

```
RESULT   4
US-09-242-690A-15
; Sequence 15, Application US/09242690A
; Patent No. 6284534
; GENERAL INFORMATION:
;  APPLICANT: KONDO, KEIJI
;  APPLICANT:  MIURA, YUTAKA
;  TITLE OF INVENTION: YEAST VECTOR AND METHOD OF PRODUCING PROTEINS USING THE
;  TITLE OF INVENTION:  SAME
;  FILE REFERENCE: 049441/0118
;  CURRENT APPLICATION NUMBER: US/09/242,690A
;  CURRENT FILING DATE:  1999-02-23
;  PRIOR APPLICATION NUMBER: PCT/JP97/02924
;  PRIOR FILING DATE: 1997-08-22
;  PRIOR APPLICATION NUMBER: JP 8/241062
;  PRIOR FILING DATE: 1996-08-23
;  NUMBER OF SEQ ID NOS: 66
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
;   LENGTH: 559
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic protein
US-09-242-690A-15
```

```
  Query Match           59.2%;  Score 1748.5;  DB 4;  Length 559;
  Best Local Similarity  58.9%;  Pred. No. 1.3e-162;
  Matches  329;  Conservative   87;  Mismatches  136;  Indels    7;  Gaps    5;

Qy       3 SFGGNIEKNKGIFKLWAPYVNSVKLK-LSKKLIPMEKNDEGFFEVEIDDIEENLTYSYII 61
           :|   |: |: || |||||  ||||| | | |  ||::::|:| : :::::    | |::
Db       2 TFAYKIDGNEVIFTLWAPYQKSVKLKVLEKGLYEMERDEKGYFTITLNNVKVRDRYKYVL 61

Qy      62 EDKREIPDPASRYQPLGVHDKSQLIRTDYQILDLGKVKIEDLIIYELHVGTFSQEGNFKG 121
           :|  |||||||||| ||| ||:|:   :  :   :| |||||||:|||||: || |:|
Db      62 DDASEIPDPASRYQPEGVHGPSQIIQESKEFNNETFLKKEDLIIYEIHVGTFTPEGTFEG 121

Qy     122 VIEKLDYLKDLGITGIELMPVAQFPGNRDWGYDGVFLYAVQNTYGGPWELAKLVNEAHKR 181
           || |||||||||| ||:||:|||| |||||||:||||||:||||    |||:||||:
Db     122 VIRKLDYLKDLGITAIEIMPIAQFPGKRDWGYDGVYLYAVQNSYGGPEGFRKLVDEAHKK 181

Qy     182 GIAVILDVVYNHIGPEGNYLLGLGPYFSDRYKTPWGLTFNFDDRGCDQVRKFILENVEYW 241
           |: |||||||||:|||||:: |||||| :|||||||||||||   |:||||||||||||
Db     182 GLGVILDVVYNHVGPEGNYMVKLGPYFSQKYKTPWGLTFNFDDAESDEVRKFILENVEYW 241
```

```
Qy    242 FKTFKIDGLRLDAVHAIFDNSPKHILQEIAEKAHQLGKFVIAESDLNDPKIV--KDDCGY 299
          | : :|| |||||||| | ||||||:|||:  |:  : ||||||||||::|  |: |||
Db    242 IKEYNVDGFRLDAVHAIIDTSPKHILEEIADVVHKYNRIVIAESDLNDPRVVNPKEKCGY 301

Qy    300 KIDAQWVDDFHHAVHAFITKEKDYYYQDFGRIEDIEKTFKDVFVYDGKYSRYRGRTHGAP 359
          ||||||||||::||::| |:  || ||| ::|| |::|||||||||| :| :||| |
Db    302 NIDAQWVDDFHHSIHAYLTGERQGYYTDFGNLDDIVKSYKDVFVYDGKYSNFRRKTHGEP 361

Qy    360 VGDLPPRKFVVFIQNHDQVGNRGNGERLSILTDKTTYLMAATLYILSPYIPLIFMGEEYY 419
          ||:|    |||:|||||||||| |||:  | |: :| :|| ||:||||||:|||||||
Db    362 VGELDGCNFVVYIQNHDQVGNRGKGERIIKLVDRESYKIAAALYLLSPYIPMIFMGEEYG 421

Qy    420 ETNPFFFFSDFSDPVLIKGVREGRLKENNQMIDPQSEEAFLKSKLSWKIDEEVLDYYKQL 479
          | |||:|||||||  ||:|||||| ||| |  ||| |  | ||||||||||:  :|| |
Db    422 EENPFYFFSDFSDSKLIQGVREGRKKENGQDTDPQDESTFNASKLSWKIDEEIFSFYKIL 481

Qy    480 INIRKRYN-NCKRVKEVRREGNCITLIMEKIGIIASFDDIVINSKITGNLLIGI--GFPK 536
          | :||  :  | |  |    | : :  :  :  |    |  | :| ||:     ||:
Db    482 IKMRKELSIACDRRVNVVNGENWLIIKGREYFSLYVFSKSSIEVKYSGTLLLSSNNSFPQ 541

Qy    537 KLKKDELIKVNRGVGVYQL 555
          ::: :  : ::|  :|:|
Db    542 HIEEGK-YEFDKGFALYKL 559


RESULT   5
US-08-399-646-2
; Sequence 2, Application US/08399646
; Patent No. 5556781
;  GENERAL INFORMATION:
;    APPLICANT:  KUBOTA, Michio
;    APPLICANT:  TSUSAKI, Keiji
;    APPLICANT:  HATTORI, Kazuko
;    APPLICANT:  SUGIMOTO, Toshiyuki
;    TITLE OF INVENTION:  DNA ENCODING ENZYME, RECOMBINANT DNA AND
;    TITLE OF INVENTION:  ENZYME, TRANSFORMANT, AND THEIR PREPARATIONS AND USE
;    NUMBER OF SEQUENCES:  18
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BROWDY AND NEIMARK
;      STREET:  419 Seventh Street, N.W., Suite 300
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/399,646
;      FILING DATE:  07-MAR-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 59834
;      FILING DATE:  07-MAR-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 59840
;      FILING DATE:  07-MAR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BROWDY, Roger L.
;      REGISTRATION NUMBER:  25,618
;      REFERENCE/DOCKET NUMBER:  KUBOTA=5
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-628-5197
;      TELEFAX:  202-737-3528
;      TELEX:  248633
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  589 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-399-646-2


  Query Match             32.8%;  Score 969;  DB 1;  Length 589;
  Best Local Similarity   39.5%;  Pred. No. 3.2e-86;
  Matches  215;  Conservative   74;  Mismatches  195;  Indels   60;  Gaps   12;

Qy     13 GIFKLWAPYVNSVKLKLSKKLIPMEKN------DEGFFEVEIDDIEENLTYSYIIE-DKR 65
          | | :|||   :| |   :  | :      |||::       :: | |::: |:
Db      8 GRFDIWAPEAGTVTLLAGGERYEMGRRPGNGPADEGWWTAADAPTGADVDYGYLLDGDEI 67

Qy     66 EIPDPASRYQPLGVHDKSQLIRTDYQILDLGKVKIEDL----------IIYELHVGTFSQ 115
          :||| :| || |||    | ||     | |  : :|          :||||||:|||:
Db     68 PLPDPRTRRQPEGVH---ALSRT----FDPGAHRWQDAGWQGRELQGSVIYELHIGTFTP 120

Qy    116 EGNFKGVIEKLDYLKDLGITGIELMPVAQFPGNRDWGYDGVFLYAVQNTYGGPWELAKLV 175
          ||       |||||  ||| |||:||  | |  :||||||  :||   ||||    : |
Db    121 EGTLDAAAGKLDYLAGLGIDFIELLPVNAFNGTHNWGYDGVQWFAVHEGYGGPAAYQRFV 180

Qy    176 NEAHKRGIAVILDVVYNHIGPEGNYLLGLGPYFSDRYKTPWGLTFNFDDRGCDQVRKFIL 235
          : ||  |: || ||||||:|| ||||   |||        || : | |  | | ||::||
Db    181 DAAHAAGLGVIQDVVYNHLGPSGNYLPRYGPYLKHGEGNTWGDSVNLDGPGSDHVRQYIL 240

Qy    236 ENVEYWFKTFKIDGLRLDAVHAIFDNSPKHILQEIAEKAHQLGK------FVIAESDLND 289
          :||  | : :::||||||||||: |    |||:|    |  |         :||||||:
Db    241 DNVAMWLRDYRVDGLRLDAVHALKDERAVHILEEFGALADALSSEGGRPLTLIAESDLNN 300

Qy    290 PKIV--KDDCGYKIDAQWVDDFHHAVHAFITKEKDYYYQDFGRIEDIEKTFKDVFVYDGK 347
          |:::  :|  || :  || |||||||  :: |   || ||  :  : |  :| | :||
Db    301 PRLLYPRDVNGYGLAGQWSDDFHHAVHVNVSGETTGYYSDFDSLGALAKVLRDGFFHDGS 360

Qy    348 YSRYRGRTHGAPV--GDLPPRKFVVFIQNHDQVGNRGNGERLSILTDKTTYLMAATLYIL 405
          || :||| || |:    : |   ||  |||||:|||  |:|||     :  :|| | :
Db    361 YSSFRGRCHGRPINFSAVHPAALVVCSQNHDQIGNRATGDRLSQSLPYGSLALAAVLTLT 420

Qy    406 SPYIPLIFMGEEYYETNPFFFFSDFSDPVLIKGVREGRLKENNQM-------IDPQSEEA 458
          |: |::|||||  | |: ||:   :| | |   |||::|  :|        ||| |
Db    421 GPFTPMLFMGEEYGATTPWQFFTSHPEPELGKATAEGRIREFERMGWDPAVVPDPQDPET 480

Qy    459 FLKSKLSWK-----IDEEVLDYYKQLINIRKRYNNCKRVKEVRREGNCITLIMEKIGIIA 513
          | :||| |          :|: |: || :|      :   |: | |   | :
Db    481 FTRSKLDWAEASAGDHARLLELYRSLITLR------RSTPELARLGFADTAV-------- 526

Qy    514 SFDD 517
            |||
Db    527 EFDD 530


RESULT   6
US-08-607-321-2
; Sequence 2, Application US/08607321
; Patent No. 5716813
;  GENERAL INFORMATION:
;    APPLICANT:  KUBOTA, Michio
;    APPLICANT:  TSUSAKI, Keiji
;    APPLICANT:  HATTORI, Kazuko
;    APPLICANT:  SUGIMOTO, Toshiyuki
;    TITLE OF INVENTION:  DNA ENCODING ENZYME, RECOMBINANT DNA AND
;    TITLE OF INVENTION:  ENZYME, TRANSFORMANT, AND THEIR PREPARATIONS AND USE
;    NUMBER OF SEQUENCES:  18
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BROWDY AND NEIMARK
;      STREET:  419 Seventh Street, N.W., Suite 300
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
```

```
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/607,321
;      FILING DATE:  26-FEB-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/399,646
;      FILING DATE:  07-MAR-1995
;      APPLICATION NUMBER:  JP 59834
;      FILING DATE:  07-MAR-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 59840
;      FILING DATE:  07-MAR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BROWDY, Roger L.
;      REGISTRATION NUMBER:  25,618
;      REFERENCE/DOCKET NUMBER:  KUBOTA=5
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-628-5197
;      TELEFAX:  202-737-3528
;      TELEX:  248633
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  589 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-607-321-2


  Query Match             32.8%;  Score 969;  DB 1;  Length 589;
  Best Local Similarity   39.5%;  Pred. No. 3.2e-86;
  Matches  215;  Conservative   74;  Mismatches  195;  Indels   60;  Gaps   12;

Qy      13 GIFKLWAPYVNSVKLKLSKKLIPMEKN------DEGFFEVEIDDIEENLTYSYIIE-DKR 65
           | | :|||   :| |    :   | :     |||::        :: | |::: |:
Db       8 GRFDIWAPEAGTVTLLAGGERYEMGRRPGNGPADEGWWTAADAPTGADVDYGYLLDGDEI 67

Qy      66 EIPDPASRYQPLGVHDKSQLIRTDYQILDLGKVKIEDL----------IIYELHVGTFSQ 115
           :||| :| || |||    | ||     | |  : :|           :|||||:|||:
Db      68 PLPDPRTRRQPEGVH---ALSRT----FDPGAHRWQDAGWQGRELQGSVIYELHIGTFTP 120

Qy     116 EGNFKGVIEKLDYLKDLGITGIELMPVAQFPGNRDWGYDGVFLYAVQNTYGGPWELAKLV 175
           ||       |||||  |||  |||:||  | |  :||||||  :||   ||||    : |
Db     121 EGTLDAAAGKLDYLAGLGIDFIELLPVNAFNGTHNWGYDGVQWFAVHEGYGGPAAYQRFV 180

Qy     176 NEAHKRGIAVILDVVYNHIGPEGNYLLGLGPYFSDRYKTPWGLTFNFDDRGCDQVRKFIL 235
           : ||  |: || ||||||:|| ||||   |||        || : | |  | | ||::||
Db     181 DAAHAAGLGVIQDVVYNHLGPSGNYLPRYGPYLKHGEGNTWGDSVNLDGPGSDHVRQYIL 240

Qy     236 ENVEYWFKTFKIDGLRLDAVHAIFDNSPKHILQEIAEKAHQLGK------FVIAESDLND 289
           :|| | : :::|||||||||||: |  |||:|    |  |         :|||||||:
Db     241 DNVAMWLRDYRVDGLRLDAVHALKDERAVHILEEFGALADALSSEGGRPLTLIAESDLNN 300

Qy     290 PKIV--KDDCGYKIDAQWVDDFHHAVHAFITKEKDYYYQDFGRIEDIEKTFKDVFVYDGK 347
           |:::  :|  || :  || |||||||  :: |   || ||  :  : |  :| | :||
Db     301 PRLLYPRDVNGYGLAGQWSDDFHHAVHVNVSGETTGYYSDFDSLGALAKVLRDGFFHDGS 360

Qy     348 YSRYRGRTHGAPV--GDLPPRKFVVFIQNHDQVGNRGNGERLSILTDKTTYLMAATLYIL 405
           || :||| || |:    : |   ||  ||||:|||  |:|||      :  :|| | :
Db     361 YSSFRGRCHGRPINFSAVHPAALVVCSQNHDQIGNRATGDRLSQSLPYGSLALAAVLTLT 420

Qy     406 SPYIPLIFMGEEYYETNPFFFFSDFSDPVLIKGVREGRLKENNQM-------IDPQSEEA 458
           |: |::|||||  | |: ||:   :| | |   |||::|  :|        |||  |
Db     421 GPFTPMLFMGEEYGATTPWQFFTSHPEPELGKATAEGRIREFERMGWDPAVVPDPQDPET 480

Qy     459 FLKSKLSWK-----IDEEVLDYYKQLINIRKRYNNCKRVKEVRREGNCITLIMEKIGIIA 513
           | :||| |          :|: |: || :|      :   |: | |   | :
Db     481 FTRSKLDWAEASAGDHARLLELYRSLITLR------RSTPELARLGFADTAV-------- 526

Qy     514 SFDD 517
             |||
Db     527 EFDD 530
```

```
RESULT   7
US-08-961-240-2
; Sequence 2, Application US/08961240
; Patent No. 5830715
;  GENERAL INFORMATION:
;    APPLICANT:  KUBOTA, Michio
;    APPLICANT:  TSUSAKI, Keiji
;    APPLICANT:  HATTORI, Kazuko
;    APPLICANT:  SUGIMOTO, Toshiyuki
;    TITLE OF INVENTION:  DNA ENCODING ENZYME, RECOMBINANT DNA AND
;    TITLE OF INVENTION:  ENZYME, TRANSFORMANT, AND THEIR PREPARATIONS AND USE
;    NUMBER OF SEQUENCES:  18
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BROWDY AND NEIMARK
;      STREET:  419 Seventh Street, N.W., Suite 300
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/961,240
;      FILING DATE:  30-OCT-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/399,646
;      FILING DATE:  07-MAR-1995
;      APPLICATION NUMBER:  JP 59834
;      FILING DATE:  07-MAR-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 59840
;      FILING DATE:  07-MAR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BROWDY, Roger L.
;      REGISTRATION NUMBER:  25,618
;      REFERENCE/DOCKET NUMBER:  KUBOTA=5
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-628-5197
;      TELEFAX:  202-737-3528
;      TELEX:  248633
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  589 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-961-240-2


  Query Match             32.8%;  Score 969;  DB 2;  Length 589;
  Best Local Similarity   39.5%;  Pred. No. 3.2e-86;
  Matches  215;  Conservative   74;  Mismatches  195;  Indels   60;  Gaps   12;

Qy      13 GIFKLWAPYVNSVKLKLSKKLIPMEKN------DEGFFEVEIDDIEENLTYSYIIE-DKR 65
           | | :|||   :| |    :   | :       |||::        :: | |::: |:
Db       8 GRFDIWAPEAGTVTLLAGGERYEMGRRPGNGPADEGWWTAADAPTGADVDYGYLLDGDEI 67

Qy      66 EIPDPASRYQPLGVHDKSQLIRTDYQILDLGKVKIEDL----------IIYELHVGTFSQ 115
           :||| :| || |||    | ||     | |  : :|           :|||||:|||:
Db      68 PLPDPRTRRQPEGVH---ALSRT----FDPGAHRWQDAGWQGRELQGSVIYELHIGTFTP 120

Qy     116 EGNFKGVIEKLDYLKDLGITGIELMPVAQFPGNRDWGYDGVFLYAVQNTYGGPWELAKLV 175
           ||       |||||  |||  |||:||  | |  :||||||  :||   ||||    : |
Db     121 EGTLDAAAGKLDYLAGLGIDFIELLPVNAFNGTHNWGYDGVQWFAVHEGYGGPAAYQRFV 180

Qy     176 NEAHKRGIAVILDVVYNHIGPEGNYLLGLGPYFSDRYKTPWGLTFNFDDRGCDQVRKFIL 235
           : ||  |: || ||||||:|| ||||   |||        || : | |  | | ||::||
Db     181 DAAHAAGLGVIQDVVYNHLGPSGNYLPRYGPYLKHGEGNTWGDSVNLDGPGSDHVRQYIL 240
```

```
Qy    236 ENVEYWFKTFKIDGLRLDAVHAIFDNSPKHILQEIAEKAHQLGK------FVIAESDLND 289
          :||  | : :::||||||||||: |    |||:|    |  |          :|||||||:
Db    241 DNVAMWLRDYRVDGLRLDAVHALKDERAVHILEEFGALADALSSEGGRPLTLIAESDLNN 300

Qy    290 PKIV--KDDCGYKIDAQWVDDFHHAVHAFITKEKDYYYQDFGRIEDIEKTFKDVFVYDGK 347
          |:::  :|  || :  || ||||||||  :: |   || ||  :  : |  :| | :||
Db    301 PRLLYPRDVNGYGLAGQWSDDFHHAVHVNVSGETTGYYSDFDSLGALAKVLRDGFFHDGS 360

Qy    348 YSRYRGRTHGAPV--GDLPPRKFVVFIQNHDQVGNRGNGERLSILTDKTTYLMAATLYIL 405
          || :||| || |:    : |   ||  |||||:|||  |:|||      :  :|| | :
Db    361 YSSFRGRCHGRPINFSAVHPAALVVCSQNHDQIGNRATGDRLSQSLPYGSLALAAVLTLT 420

Qy    406 SPYIPLIFMGEEYYETNPFFFFSDFSDPVLIKGVREGRLKENNQM-------IDPQSEEA 458
          |: |::|||||  | |: ||:   :| | |   |||::|  :|        |||  |
Db    421 GPFTPMLFMGEEYGATTPWQFFTSHPEPELGKATAEGRIREFERMGWDPAVVPDPQDPET 480

Qy    459 FLKSKLSWK-----IDEEVLDYYKQLINIRKRYNNCKRVKEVRREGNCITLIMEKIGIIA 513
          | :||| |          :|: |: || :|      :   |: | |   | :
Db    481 FTRSKLDWAEASAGDHARLLELYRSLITLR------RSTPELARLGFADTAV-------- 526

Qy    514 SFDD 517
           |||
Db    527 EFDD 530


RESULT   8
US-08-605-501-2
; Sequence 2, Application US/08605501
; Patent No. 5834287
;  GENERAL INFORMATION:
;    APPLICANT:  KUBOTA, Michio
;    APPLICANT:  TSUSAKI, Keiji
;    APPLICANT:  HATTORI, Kazuko
;    APPLICANT:  SUGIMOTO, Toshiyuki
;    TITLE OF INVENTION:  DNA ENCODING ENZYME, RECOMBINANT DNA AND
;    TITLE OF INVENTION:  ENZYME, TRANSFORMANT, AND THEIR PREPARATIONS AND USE
;    NUMBER OF SEQUENCES:  18
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BROWDY AND NEIMARK
;      STREET:  419 Seventh Street, N.W., Suite 300
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/605,501
;      FILING DATE:  26-FEB-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/399,646
;      FILING DATE:  07-MAR-1995
;      APPLICATION NUMBER:  JP 59834
;      FILING DATE:  07-MAR-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 59840
;      FILING DATE:  07-MAR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BROWDY, Roger L.
;      REGISTRATION NUMBER:  25,618
;      REFERENCE/DOCKET NUMBER:  KUBOTA=5
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-628-5197
;      TELEFAX:  202-737-3528
;      TELEX:  248633
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  589 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-605-501-2


  Query Match            32.8%;  Score 969;  DB 2;  Length 589;
  Best Local Similarity  39.5%;  Pred. No. 3.2e-86;
  Matches  215;  Conservative   74;  Mismatches  195;  Indels   60;  Gaps   12;

Qy     13 GIFKLWAPYVNSVKLKLSKKLIPMEKN------DEGFFEVEIDDIEENLTYSYIIE-DKR 65
          | | :|||   :| |    :   | :       |||::          :: | |::: |:
Db      8 GRFDIWAPEAGTVTLLAGGERYEMGRRPGNGPADEGWWTAADAPTGADVDYGYLLDGDEI 67

Qy     66 EIPDPASRYQPLGVHDKSQLIRTDYQILDLGKVKIEDL----------IIYELHVGTFSQ 115
           :||| :| || |||    | ||     | |  : :|           :|||||:|||:
Db     68 PLPDPRTRRQPEGVH---ALSRT----FDPGAHRWQDAGWQGRELQGSVIYELHIGTFTP 120

Qy    116 EGNFKGVIEKLDYLKDLGITGIELMPVAQFPGNRDWGYDGVFLYAVQNTYGGPWELAKLV 175
          ||       |||||  |||  |||:||  | |  :||||||  :||   ||||    : |
Db    121 EGTLDAAAGKLDYLAGLGIDFIELLPVNAFNGTHNWGYDGVQWFAVHEGYGGPAAYQRFV 180

Qy    176 NEAHKRGIAVILDVVYNHIGPEGNYLLGLGPYFSDRYKTPWGLTFNFDDRGCDQVRKFIL 235
          : ||  |: || ||||||:|| ||||   |||        || : | |  | | ||::||
Db    181 DAAHAAGLGVIQDVVYNHLGPSGNYLPRYGPYLKHGEGNTWGDSVNLDGPGSDHVRQYIL 240

Qy    236 ENVEYWFKTFKIDGLRLDAVHAIFDNSPKHILQEIAEKAHQLGK------FVIAESDLND 289
          :||  | : :::||||||||||: |    |||:|    |  |          :|||||||:
Db    241 DNVAMWLRDYRVDGLRLDAVHALKDERAVHILEEFGALADALSSEGGRPLTLIAESDLNN 300

Qy    290 PKIV--KDDCGYKIDAQWVDDFHHAVHAFITKEKDYYYQDFGRIEDIEKTFKDVFVYDGK 347
          |:::  :|  || :  || ||||||||  :: |   || ||  :  : |  :| | :||
Db    301 PRLLYPRDVNGYGLAGQWSDDFHHAVHVNVSGETTGYYSDFDSLGALAKVLRDGFFHDGS 360

Qy    348 YSRYRGRTHGAPV--GDLPPRKFVVFIQNHDQVGNRGNGERLSILTDKTTYLMAATLYIL 405
          || :||| || |:    : |   ||  |||||:|||  |:|||      :  :|| | :
Db    361 YSSFRGRCHGRPINFSAVHPAALVVCSQNHDQIGNRATGDRLSQSLPYGSLALAAVLTLT 420

Qy    406 SPYIPLIFMGEEYYETNPFFFFSDFSDPVLIKGVREGRLKENNQM-------IDPQSEEA 458
          |: |::|||||  | |: ||:   :| | |   |||::|  :|        |||  |
Db    421 GPFTPMLFMGEEYGATTPWQFFTSHPEPELGKATAEGRIREFERMGWDPAVVPDPQDPET 480

Qy    459 FLKSKLSWK-----IDEEVLDYYKQLINIRKRYNNCKRVKEVRREGNCITLIMEKIGIIA 513
          | :||| |          :|: |: || :|      :   |: | |   | :
Db    481 FTRSKLDWAEASAGDHARLLELYRSLITLR------RSTPELARLGFADTAV-------- 526

Qy    514 SFDD 517
           |||
Db    527 EFDD 530


RESULT   9
US-08-399-646-12
; Sequence 12, Application US/08399646
; Patent No. 5556781
;  GENERAL INFORMATION:
;    APPLICANT:  KUBOTA, Michio
;    APPLICANT:  TSUSAKI, Keiji
;    APPLICANT:  HATTORI, Kazuko
;    APPLICANT:  SUGIMOTO, Toshiyuki
;    TITLE OF INVENTION:  DNA ENCODING ENZYME, RECOMBINANT DNA AND
;    TITLE OF INVENTION:  ENZYME, TRANSFORMANT, AND THEIR PREPARATIONS AND USE
;    NUMBER OF SEQUENCES:  18
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BROWDY AND NEIMARK
;      STREET:  419 Seventh Street, N.W., Suite 300
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
```

```
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/399,646
;      FILING DATE:  07-MAR-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 59834
;      FILING DATE:  07-MAR-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 59840
;      FILING DATE:  07-MAR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BROWDY, Roger L.
;      REGISTRATION NUMBER:  25,618
;      REFERENCE/DOCKET NUMBER:  KUBOTA=5
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-628-5197
;      TELEFAX:  202-737-3528
;      TELEX:  248633
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  596 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-399-646-12


  Query Match             32.8%;  Score 969;  DB 1;  Length 596;
  Best Local Similarity   39.5%;  Pred. No. 3.2e-86;
  Matches  215;  Conservative   74;  Mismatches  195;  Indels   60;  Gaps   12;

Qy      13 GIFKLWAPYVNSVKLKLSKKLIPMEKN------DEGFFEVEIDDIEENLTYSYIIE-DKR 65
           | | :|||   :| |    :   | :       |||::         :: | |::: |:
Db      15 GRFDIWAPEAGTVTLLAGGERYEMGRRPGNGPADEGWWTAADAPTGADVDYGYLLDGDEI 74

Qy      66 EIPDPASRYQPLGVHDKSQLIRTDYQILDLGKVKIEDL----------IIYELHVGTFSQ 115
             :||| :| || |||    | ||     | |  : :|           :|||||:|||:
Db      75 PLPDPRTRRQPEGVH---ALSRT----FDPGAHRWQDAGWQGRELQGSVIYELHIGTFTP 127

Qy     116 EGNFKGVIEKLDYLKDLGITGIELMPVAQFPGNRDWGYDGVFLYAVQNTYGGPWELAKLV 175
           ||       |||||  |||  |||:||  | |  :||||||  :||   ||||    : |
Db     128 EGTLDAAAGKLDYLAGLGIDFIELLPVNAFNGTHNWGYDGVQWFAVHEGYGGPAAYQRFV 187

Qy     176 NEAHKRGIAVILDVVYNHIGPEGNYLLGLGPYFSDRYKTPWGLTFNFDDRGCDQVRKFIL 235
           : ||  |: || ||||||:|| ||||   |||        || : | |  | | ||::||
Db     188 DAAHAAGLGVIQDVVYNHLGPSGNYLPRYGPYLKHGEGNTWGDSVNLDGPGSDHVRQYIL 247

Qy     236 ENVEYWFKTFKIDGLRLDAVHAIFDNSPKHILQEIAEKAHQLGK------FVIAESDLND 289
           :||  | : :::|||||||||: |    |||:|    |  |         :|||||||:
Db     248 DNVAMWLRDYRVDGLRLDAVHALKDERAVHILEEFGALADALSSEGGRPLTLIAESDLNN 307

Qy     290 PKIV--KDDCGYKIDAQWVDDFHHAVHAFITKEKDYYYQDFGRIEDIEKTFKDVFVYDGK 347
           |:::  :|  || :  || ||||||||  :: |   || ||  :  : |  :| | :||
Db     308 PRLLYPRDVNGYGLAGQWSDDFHHAVHVNVSGETTGYYSDFDSLGALAKVLRDGFFHDGS 367

Qy     348 YSRYRGRTHGAPV--GDLPPRKFVVFIQNHDQVGNRGNGERLSILTDKTTYLMAATLYIL 405
           || :||| || |:    : |   ||  |||||:|||  |:|||      :  :|| | :
Db     368 YSSFRGRCHGRPINFSAVHPAALVVCSQNHDQIGNRATGDRLSQSLPYGSLALAAVLTLT 427

Qy     406 SPYIPLIFMGEEYYETNPFFFFSDFSDPVLIKGVREGRLKENNQM-------IDPQSEEA 458
             |: |::|||||  | |: ||:   :| | |   |||::|  :|        |||  |
Db     428 GPFTPMLFMGEEYGATTPWQFFTSHPEPELGKATAEGRIREFERMGWDPAVVPDPQDPET 487

Qy     459 FLKSKLSWK-----IDEEVLDYYKQLINIRKRYNNCKRVKEVRREGNCITLIMEKIGIIA 513
           | :||| |          :|: |: || :|      :   |: | |   | :
Db     488 FTRSKLDWAEASAGDHARLLELYRSLITLR------RSTPELARLGFADTAV-------- 533

Qy     514 SFDD 517
             |||
Db     534 EFDD 537
```

```
RESULT  10
US-08-607-321-12
; Sequence 12, Application US/08607321
; Patent No. 5716813
;  GENERAL INFORMATION:
;    APPLICANT:  KUBOTA, Michio
;    APPLICANT:  TSUSAKI, Keiji
;    APPLICANT:  HATTORI, Kazuko
;    APPLICANT:  SUGIMOTO, Toshiyuki
;    TITLE OF INVENTION:  DNA ENCODING ENZYME, RECOMBINANT DNA AND
;    TITLE OF INVENTION:  ENZYME, TRANSFORMANT, AND THEIR PREPARATIONS AND USE
;    NUMBER OF SEQUENCES:  18
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BROWDY AND NEIMARK
;      STREET:  419 Seventh Street, N.W., Suite 300
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/607,321
;      FILING DATE:  26-FEB-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/399,646
;      FILING DATE:  07-MAR-1995
;      APPLICATION NUMBER:  JP 59834
;      FILING DATE:  07-MAR-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 59840
;      FILING DATE:  07-MAR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BROWDY, Roger L.
;      REGISTRATION NUMBER:  25,618
;      REFERENCE/DOCKET NUMBER:  KUBOTA=5
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-628-5197
;      TELEFAX:  202-737-3528
;      TELEX:  248633
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  596 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-607-321-12


  Query Match             32.8%;  Score 969;  DB 1;  Length 596;
  Best Local Similarity   39.5%;  Pred. No. 3.2e-86;
  Matches  215;  Conservative   74;  Mismatches  195;  Indels   60;  Gaps   12;

Qy      13 GIFKLWAPYVNSVKLKLSKKLIPMEKN------DEGFFEVEIDDIEENLTYSYIIE-DKR 65
           | | :|||   :| |    :   | :       |||::         :: | |::: |:
Db      15 GRFDIWAPEAGTVTLLAGGERYEMGRRPGNGPADEGWWTAADAPTGADVDYGYLLDGDEI 74

Qy      66 EIPDPASRYQPLGVHDKSQLIRTDYQILDLGKVKIEDL----------IIYELHVGTFSQ 115
             :||| :| || |||    | ||     | |  : :|           :|||||:|||:
Db      75 PLPDPRTRRQPEGVH---ALSRT----FDPGAHRWQDAGWQGRELQGSVIYELHIGTFTP 127

Qy     116 EGNFKGVIEKLDYLKDLGITGIELMPVAQFPGNRDWGYDGVFLYAVQNTYGGPWELAKLV 175
           ||       |||||  |||  |||:||  | |  :||||||  :||   ||||    : |
Db     128 EGTLDAAAGKLDYLAGLGIDFIELLPVNAFNGTHNWGYDGVQWFAVHEGYGGPAAYQRFV 187

Qy     176 NEAHKRGIAVILDVVYNHIGPEGNYLLGLGPYFSDRYKTPWGLTFNFDDRGCDQVRKFIL 235
           : ||  |: || ||||||:|| ||||   |||        || : | |  | | ||::||
Db     188 DAAHAAGLGVIQDVVYNHLGPSGNYLPRYGPYLKHGEGNTWGDSVNLDGPGSDHVRQYIL 247
```

```
Qy    236 ENVEYWFKTFKIDGLRLDAVHAIFDNSPKHILQEIAEKAHQLGK------FVIAESDLND 289
          :||  | : :::|||||||||||: |    |||:|    |  |         :|||||||:
Db    248 DNVAMWLRDYRVDGLRLDAVHALKDERAVHILEEFGALADALSSEGGRPLTLIAESDLNN 307

Qy    290 PKIV--KDDCGYKIDAQWVDDFHHAVHAFITKEKDYYYQDFGRIEDIEKTFKDVFVYDGK 347
          |:::  :|  || :  || ||||||||  :: |   || ||  :  : |  :| | :||
Db    308 PRLLYPRDVNGYGLAGQWSDDFHHAVHVNVSGETTGYYSDFDSLGALAKVLRDGFFHDGS 367

Qy    348 YSRYRGRTHGAPV--GDLPPRKFVVFIQNHDQVGNRGNGERLSILTDKTTYLMAATLYIL 405
          || :||| || |:    : |   ||  |||||:|||  |:|||      :  :|| | :
Db    368 YSSFRGRCHGRPINFSAVHPAALVVCSQNHDQIGNRATGDRLSQSLPYGSLALAAVLTLT 427

Qy    406 SPYIPLIFMGEEYYETNPFFFFSDFSDPVLIKGVREGRLKENNQM-------IDPQSEEA 458
           |: |::|||||  | |: ||:   :| | |   |||::|  :|        |||  |
Db    428 GPFTPMLFMGEEYGATTPWQFFTSHPEPELGKATAEGRIREFERMGWDPAVVPDPQDPET 487

Qy    459 FLKSKLSWK-----IDEEVLDYYKQLINIRKRYNNCKRVKEVRREGNCITLIMEKIGIIA 513
          | :||| |          :|: |: || :|      :   |: | |   | :
Db    488 FTRSKLDWAEASAGDHARLLELYRSLITLR------RSTPELARLGFADTAV-------- 533

Qy    514 SFDD 517
           |||
Db    534 EFDD 537


RESULT  11
US-08-961-240-12
; Sequence 12, Application US/08961240
; Patent No. 5830715
;  GENERAL INFORMATION:
;    APPLICANT:  KUBOTA, Michio
;    APPLICANT:  TSUSAKI, Keiji
;    APPLICANT:  HATTORI, Kazuko
;    APPLICANT:  SUGIMOTO, Toshiyuki
;    TITLE OF INVENTION:  DNA ENCODING ENZYME, RECOMBINANT DNA AND
;    TITLE OF INVENTION:  ENZYME, TRANSFORMANT, AND THEIR PREPARATIONS AND USE
;    NUMBER OF SEQUENCES:  18
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BROWDY AND NEIMARK
;      STREET:  419 Seventh Street, N.W., Suite 300
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/961,240
;      FILING DATE:  30-OCT-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/399,646
;      FILING DATE:  07-MAR-1995
;      APPLICATION NUMBER:  JP 59834
;      FILING DATE:  07-MAR-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 59840
;      FILING DATE:  07-MAR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BROWDY, Roger L.
;      REGISTRATION NUMBER:  25,618
;      REFERENCE/DOCKET NUMBER:  KUBOTA=5
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-628-5197
;      TELEFAX:  202-737-3528
;      TELEX:  248633
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  596 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-961-240-12


  Query Match             32.8%;  Score 969;  DB 2;  Length 596;
  Best Local Similarity   39.5%;  Pred. No. 3.2e-86;
  Matches  215;  Conservative   74;  Mismatches  195;  Indels   60;  Gaps   12;

Qy     13 GIFKLWAPYVNSVKLKLSKKLIPMEKN------DEGFFEVEIDDIEENLTYSYIIE-DKR 65
          | | :|||   :| |    :   | :       |||::          :: | |::: |:
Db     15 GRFDIWAPEAGTVTLLAGGERYEMGRRPGNGPADEGWWTAADAPTGADVDYGYLLDGDEI 74

Qy     66 EIPDPASRYQPLGVHDKSQLIRTDYQILDLGKVKIEDL----------IIYELHVGTFSQ 115
           :||| :| || |||     | ||     | |  : :|           :|||||:|||:
Db     75 PLPDPRTRRQPEGVH---ALSRT----FDPGAHRWQDAGWQGRELQGSVIYELHIGTFTP 127

Qy    116 EGNFKGVIEKLDYLKDLGITGIELMPVAQFPGNRDWGYDGVFLYAVQNTYGGPWELAKLV 175
          ||       |||||  |||  |||:||  | |  :||||||  :||   ||||    : |
Db    128 EGTLDAAAGKLDYLAGLGIDFIELLPVNAFNGTHNWGYDGVQWFAVHEGYGGPAAYQRFV 187

Qy    176 NEAHKRGIAVILDVVYNHIGPEGNYLLGLGPYFSDRYKTPWGLTFNFDDRGCDQVRKFIL 235
          : ||  |: || ||||||:|| ||||   |||         || : | |  | | ||::||
Db    188 DAAHAAGLGVIQDVVYNHLGPSGNYLPRYGPYLKHGEGNTWGDSVNLDGPGSDHVRQYIL 247

Qy    236 ENVEYWFKTFKIDGLRLDAVHAIFDNSPKHILQEIAEKAHQLGK------FVIAESDLND 289
          :||  | : :::|||||||||||: |    |||:|    |  |         :|||||||:
Db    248 DNVAMWLRDYRVDGLRLDAVHALKDERAVHILEEFGALADALSSEGGRPLTLIAESDLNN 307

Qy    290 PKIV--KDDCGYKIDAQWVDDFHHAVHAFITKEKDYYYQDFGRIEDIEKTFKDVFVYDGK 347
          |:::  :|  || :  || ||||||||  :: |   || ||  :  : |  :| | :||
Db    308 PRLLYPRDVNGYGLAGQWSDDFHHAVHVNVSGETTGYYSDFDSLGALAKVLRDGFFHDGS 367

Qy    348 YSRYRGRTHGAPV--GDLPPRKFVVFIQNHDQVGNRGNGERLSILTDKTTYLMAATLYIL 405
          || :||| || |:    : |   ||  |||||:|||  |:|||      :  :|| | :
Db    368 YSSFRGRCHGRPINFSAVHPAALVVCSQNHDQIGNRATGDRLSQSLPYGSLALAAVLTLT 427

Qy    406 SPYIPLIFMGEEYYETNPFFFFSDFSDPVLIKGVREGRLKENNQM-------IDPQSEEA 458
           |: |::|||||  | |: ||:   :| | |   |||::|  :|        |||  |
Db    428 GPFTPMLFMGEEYGATTPWQFFTSHPEPELGKATAEGRIREFERMGWDPAVVPDPQDPET 487

Qy    459 FLKSKLSWK-----IDEEVLDYYKQLINIRKRYNNCKRVKEVRREGNCITLIMEKIGIIA 513
          | :||| |          :|: |: || :|      :   |: | |   | :
Db    488 FTRSKLDWAEASAGDHARLLELYRSLITLR------RSTPELARLGFADTAV-------- 533

Qy    514 SFDD 517
           |||
Db    534 EFDD 537


RESULT  12
US-08-605-501-12
; Sequence 12, Application US/08605501
; Patent No. 5834287
;  GENERAL INFORMATION:
;    APPLICANT:  KUBOTA, Michio
;    APPLICANT:  TSUSAKI, Keiji
;    APPLICANT:  HATTORI, Kazuko
;    APPLICANT:  SUGIMOTO, Toshiyuki
;    TITLE OF INVENTION:  DNA ENCODING ENZYME, RECOMBINANT DNA AND
;    TITLE OF INVENTION:  ENZYME, TRANSFORMANT, AND THEIR PREPARATIONS AND USE
;    NUMBER OF SEQUENCES:  18
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BROWDY AND NEIMARK
;      STREET:  419 Seventh Street, N.W., Suite 300
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
```

```
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/605,501
;       FILING DATE:  26-FEB-1996
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/399,646
;       FILING DATE:  07-MAR-1995
;       APPLICATION NUMBER:  JP 59834
;       FILING DATE:  07-MAR-1994
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  JP 59840
;       FILING DATE:  07-MAR-1994
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  BROWDY, Roger L.
;       REGISTRATION NUMBER:  25,618
;       REFERENCE/DOCKET NUMBER:  KUBOTA=5
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  202-628-5197
;       TELEFAX:  202-737-3528
;       TELEX:  248633
;   INFORMATION FOR SEQ ID NO:  12:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  596 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-605-501-12


  Query Match             32.8%;  Score 969;  DB 2;  Length 596;
  Best Local Similarity   39.5%;  Pred. No. 3.2e-86;
  Matches  215;  Conservative   74;  Mismatches  195;  Indels   60;  Gaps   12;

Qy      13 GIFKLWAPYVNSVKLKLSKKLIPMEKN------DEGFFEVEIDDIEENLTYSYIIE-DKR 65
           | | :|||   :| |    :   | :       |||::         :: | |::: |:
Db      15 GRFDIWAPEAGTVTLLAGGERYEMGRRPGNGPADEGWWTAADAPTGADVDYGYLLDGDEI 74

Qy      66 EIPDPASRYQPLGVHDKSQLIRTDYQILDLGKVKIEDL----------IIYELHVGTFSQ 115
            :||| :| || |||    | ||     | |  : :|           :|||||:|||:
Db      75 PLPDPRTRRQPEGVH---ALSRT----FDPGAHRWQDAGWQGRELQGSVIYELHIGTFTP 127

Qy     116 EGNFKGVIEKLDYLKDLGITGIELMPVAQFPGNRDWGYDGVFLYAVQNTYGGPWELAKLV 175
           ||       |||||  |||  |||:||  | |  :||||||  :||   ||||    : |
Db     128 EGTLDAAAGKLDYLAGLGIDFIELLPVNAFNGTHNWGYDGVQWFAVHEGYGGPAAYQRFV 187

Qy     176 NEAHKRGIAVILDVVYNHIGPEGNYLLGLGPYFSDRYKTPWGLTFNFDDRGCDQVRKFIL 235
           : ||  |: || ||||||:|| ||||   |||        || : | |  | | ||::||
Db     188 DAAHAAGLGVIQDVVYNHLGPSGNYLPRYGPYLKHGEGNTWGDSVNLDGPGSDHVRQYIL 247

Qy     236 ENVEYWFKTFKIDGLRLDAVHAIFDNSPKHILQEIAEKAHQLGK------FVIAESDLND 289
           :||  | : :::||||||||||: |    |||:|    |  |         :|||||||:
Db     248 DNVAMWLRDYRVDGLRLDAVHALKDERAVHILEEFGALADALSSEGGRPLTLIAESDLNN 307

Qy     290 PKIV--KDDCGYKIDAQWVDDFHHAVHAFITKEKDYYYQDFGRIEDIEKTFKDVFVYDGK 347
           |:::  :|  || :  || ||||||||  :: |   || ||  :  : | :| | :||
Db     308 PRLLYPRDVNGYGLAGQWSDDFHHAVHVNVSGETTGYYSDFDSLGALAKVLRDGFFHDGS 367

Qy     348 YSRYRGRTHGAPV--GDLPPRKFVVFIQNHDQVGNRGNGERLSILTDKTTYLMAATLYIL 405
           || :||| || |:    : |   ||  |||||:|||  |:|||      :  :|| | :
Db     368 YSSFRGRCHGRPINFSAVHPAALVVCSQNHDQIGNRATGDRLSQSLPYGSLALAAVLTLT 427

Qy     406 SPYIPLIFMGEEYYETNPFFFFSDFSDPVLIKGVREGRLKENNQM-------IDPQSEEA 458
            |: |::|||||  | |: ||:   :| | |   |||::|  :|        |||  |
Db     428 GPFTPMLFMGEEYGATTPWQFFTSHPEPELGKATAEGRIREFERMGWDPAVVPDPQDPET 487

Qy     459 FLKSKLSWK-----IDEEVLDYYKQLINIRKRYNNCKRVKEVRREGNCITLIMEKIGIIA 513
           | :||| |          :|: |: || :|      :   |: | |   | :
Db     488 FTRSKLDWAEASAGDHARLLELYRSLITLR------RSTPELARLGFADTAV-------- 533

Qy     514 SFDD 517
           |||
Db     534 EFDD 537


RESULT  13
US-08-399-646-4
; Sequence 4, Application US/08399646
; Patent No. 5556781
;  GENERAL INFORMATION:
;    APPLICANT:  KUBOTA, Michio
;    APPLICANT:  TSUSAKI, Keiji
;    APPLICANT:  HATTORI, Kazuko
;    APPLICANT:  SUGIMOTO, Toshiyuki
;    TITLE OF INVENTION:  DNA ENCODING ENZYME, RECOMBINANT DNA AND
;    TITLE OF INVENTION:  ENZYME, TRANSFORMANT, AND THEIR PREPARATIONS AND USE
;    NUMBER OF SEQUENCES:  18
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BROWDY AND NEIMARK
;      STREET:  419 Seventh Street, N.W., Suite 300
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/399,646
;      FILING DATE:  07-MAR-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 59834
;      FILING DATE:  07-MAR-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 59840
;      FILING DATE:  07-MAR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BROWDY, Roger L.
;      REGISTRATION NUMBER:  25,618
;      REFERENCE/DOCKET NUMBER:  KUBOTA=5
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-628-5197
;      TELEFAX:  202-737-3528
;      TELEX:  248633
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  597 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-399-646-4


  Query Match             32.5%;  Score 959;  DB 1;  Length 597;
  Best Local Similarity   40.2%;  Pred. No. 3.1e-85;
  Matches  202;  Conservative   81;  Mismatches  186;  Indels   34;  Gaps   10;

Qy      15 FKLWAPYVNSVKLKLSKKLIPMEK------NDEGFFEVEIDDIEENLTYSYIIE-DKREI 67
           : :|||   || |    :   |::       | |::       : |: | |::: |:  :
Db      18 YDVWAPNAESVTLLAGGERYAMQRRAETGPEDAGWWTAAGAPTDGNVDYGYLLDGDETPL 77

Qy      68 PDPASRYQPLGVHDKSQLI-RTDYQILD---LGKVKIEDLIIYELHVGTFSQEGNFKGVI 123
           ||| :| || |||  |:    : |   |    |: ::: :|||||:|||: ||  :
Db      78 PDPRTRRQPDGVHALSRTFDPSAYSWQDDAWQGR-ELQGAVIYELHLGTFTPEGTLEAAA 136

Qy     124 EKLDYLKDLGITGIELMPVAQFPGNRDWGYDGVFLYAVQNTYGGPWELAKLVNEAHKRGI 183
           |||||  ||:  |||:||  | |  :||||||  :||   ||||    : |: ||  |:
Db     137 GKLDYLAGLGVDFIELLPVNAFNGTHNWGYDGVQWFAVHEAYGGPEAYQRFVDAAHAAGL 196

Qy     184 AVILDVVYNHIGPEGNYLLGLGPYFSDRYKTPWGLTFNFDDRGCDQVRKFILENVEYWFK 243
           || ||||||:|| ||||   |||        || : | |  | | ||::||:|:  | :
Db     197 GVIQDVVYNHLGPSGNYLPRFGPYLKQGEGNTWGDSVNLDGPGSDHVRRYILDNLAMWLR 256
```

```
Qy    244 TFKIDGLRLDAVHAIFDNSPKHILQEIAEKAHQLGKFV------IAESDLNDPKIV--KD 295
          :::||||||||||: |    |||::    | |:   |      |||||||:|:::  :|
Db    257 DYRVDGLRLDAVHALKDERAVHILEDFGALADQISAEVGRPLTLIAESDLNNPRLLYPRD 316

Qy    296 DCGYKIDAQWVDDFHHAVHAFITKEKDYYYQDFGRIEDIEKTFKDVFVYDGKYSRYRGRT 355
            || :: || |||||||| :| |   || ||  :  : |  :| | :|| || :| |
Db    317 VNGYGLEGQWSDDFHHAVHVNVTGETTGYYSDFDSLAALAKVLRDGFFHDGSYSSFRERH 376

Qy    356 HGAPV--GDLPPRKFVVFIQNHDQVGNRGNGERLSILTDKTTYLMAATLYILSPYIPLIF 413
          || |:    : |   ||  |||||:|||  |:|||      :  :|| | :  |: |::
Db    377 HGRPINFSAVHPAALVVCSQNHDQIGNRATGDRLSQTLPYGSLALAAVLTLTGPFTPMLL 436

Qy    414 MGEEYYETNPFFFFSDFSDPVLIKGVREGRLKENNQM-------IDPQSEEAFLKSKLSW 466
          |||||  : |: ||:   :| | |   |||:||  :|       |||  | | :||| |
Db    437 MGEEYGASTPWQFFTSHPEPELGKATAEGRIKEFERMGWDPAVVPDPQDPETFRRSKLDW 496

Qy    467 KIDEE-----VLDYYKQLINIRK 484
            |     :|: |: |  :|:
Db    497 AEAAEGDHARLLELYRSLTALRR 519


RESULT  14
US-08-607-321-4
; Sequence 4, Application US/08607321
; Patent No. 5716813
;  GENERAL INFORMATION:
;    APPLICANT:  KUBOTA, Michio
;    APPLICANT:  TSUSAKI, Keiji
;    APPLICANT:  HATTORI, Kazuko
;    APPLICANT:  SUGIMOTO, Toshiyuki
;    TITLE OF INVENTION:  DNA ENCODING ENZYME, RECOMBINANT DNA AND
;    TITLE OF INVENTION:  ENZYME, TRANSFORMANT, AND THEIR PREPARATIONS AND USE
;    NUMBER OF SEQUENCES:  18
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BROWDY AND NEIMARK
;      STREET:  419 Seventh Street, N.W., Suite 300
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/607,321
;      FILING DATE:  26-FEB-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/399,646
;      FILING DATE:  07-MAR-1995
;      APPLICATION NUMBER:  JP 59834
;      FILING DATE:  07-MAR-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 59840
;      FILING DATE:  07-MAR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BROWDY, Roger L.
;      REGISTRATION NUMBER:  25,618
;      REFERENCE/DOCKET NUMBER:  KUBOTA=5
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-628-5197
;      TELEFAX:  202-737-3528
;      TELEX:  248633
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  597 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-607-321-4

  Query Match           32.5%;  Score 959;  DB 1;  Length 597;
  Best Local Similarity   40.2%;  Pred. No. 3.1e-85;
  Matches  202;  Conservative   81;  Mismatches  186;  Indels   34;  Gaps   10;

Qy     15 FKLWAPYVNSVKLKLSKKLIPMEK------NDEGFFEVEIDDIEENLTYSYIIE-DKREI 67
          : :|||   || |    :   |::       | |::       : |: | |::: |:  :
Db     18 YDVWAPNAESVTLLAGGERYAMQRRAETGPEDAGWWTAAGAPTDGNVDYGYLLDGDETPL 77

Qy     68 PDPASRYQPLGVHDKSQLI-RTDYQILD---LGKVKIEDLIIYELHVGTFSQEGNFKGVI 123
          ||| :| || |||  |:    : |   |    |: :::  :|||||:|||: ||  :
Db     78 PDPRTRRQPDGVHALSRTFDPSAYSWQDDAWQGR-ELQGAVIYELHLGTFTPEGTLEAAA 136

Qy    124 EKLDYLKDLGITGIELMPVAQFPGNRDWGYDGVFLYAVQNTYGGPWELAKLVNEAHKRGI 183
          |||||  ||:  |||:||  | |  :||||||  :||   ||||    : |: ||  |:
Db    137 GKLDYLAGLGVDFIELLPVNAFNGTHNWGYDGVQWFAVHEAYGGPEAYQRFVDAAHAAGL 196

Qy    184 AVILDVVYNHIGPEGNYLLGLGPYFSDRYKTPWGLTFNFDDRGCDQVRKFILENVEYWFK 243
          || |||||||:|| ||||   |||         || : | |  | | ||::||:|:  | :
Db    197 GVIQDVVYNHLGPSGNYLPRFGPYLKQGEGNTWGDSVNLDGPGSDHVRRYILDNLAMWLR 256

Qy    244 TFKIDGLRLDAVHAIFDNSPKHILQEIAEKAHQLGKFV------IAESDLNDPKIV--KD 295
          :::||||||||||: |    |||::    | |:   |      |||||||:|:::  :|
Db    257 DYRVDGLRLDAVHALKDERAVHILEDFGALADQISAEVGRPLTLIAESDLNNPRLLYPRD 316

Qy    296 DCGYKIDAQWVDDFHHAVHAFITKEKDYYYQDFGRIEDIEKTFKDVFVYDGKYSRYRGRT 355
            || :: || |||||||| :| |   || ||  :  : |  :| | :|| || :| |
Db    317 VNGYGLEGQWSDDFHHAVHVNVTGETTGYYSDFDSLAALAKVLRDGFFHDGSYSSFRERH 376

Qy    356 HGAPV--GDLPPRKFVVFIQNHDQVGNRGNGERLSILTDKTTYLMAATLYILSPYIPLIF 413
          || |:    : |   ||  |||||:|||  |:|||      :  :|| | :  |: |::
Db    377 HGRPINFSAVHPAALVVCSQNHDQIGNRATGDRLSQTLPYGSLALAAVLTLTGPFTPMLL 436

Qy    414 MGEEYYETNPFFFFSDFSDPVLIKGVREGRLKENNQM-------IDPQSEEAFLKSKLSW 466
          |||||  : |: ||:   :| | |   |||:||  :|       |||  | | :||| |
Db    437 MGEEYGASTPWQFFTSHPEPELGKATAEGRIKEFERMGWDPAVVPDPQDPETFRRSKLDW 496

Qy    467 KIDEE-----VLDYYKQLINIRK 484
            |     :|: |: |  :|:
Db    497 AEAAEGDHARLLELYRSLTALRR 519


RESULT  15
US-08-961-240-4
; Sequence 4, Application US/08961240
; Patent No. 5830715
;  GENERAL INFORMATION:
;    APPLICANT:  KUBOTA, Michio
;    APPLICANT:  TSUSAKI, Keiji
;    APPLICANT:  HATTORI, Kazuko
;    APPLICANT:  SUGIMOTO, Toshiyuki
;    TITLE OF INVENTION:  DNA ENCODING ENZYME, RECOMBINANT DNA AND
;    TITLE OF INVENTION:  ENZYME, TRANSFORMANT, AND THEIR PREPARATIONS AND USE
;    NUMBER OF SEQUENCES:  18
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BROWDY AND NEIMARK
;      STREET:  419 Seventh Street, N.W., Suite 300
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/961,240
;      FILING DATE:  30-OCT-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/399,646
```

; FILING DATE: 07-MAR-1995
; APPLICATION NUMBER: JP 59834
; FILING DATE: 07-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 59840
; FILING DATE: 07-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWDY, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: KUBOTA=5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; TELEX: 248633
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 597 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-961-240-4

Query Match 32.5%; Score 959; DB 2; Length 597;
Best Local Similarity 40.2%; Pred. No. 3.1e-85;
Matches 202; Conservative 81; Mismatches 186; Indels 34; Gaps 10;

```
Qy      15 FKLWAPYVNSVKLKLSKKLIPMEK------NDEGFFEVEIDDIEENLTYSYIIE-DKREI 67
           .  : :|||   || |    :   |::       | |::       : |: | |::: |:  :
Db      18 YDVWAPNAESVTLLAGGERYAMQRRAETGPEDAGWWTAAGAPTDGNVDYGYLLDGDETPL 77

Qy      68 PDPASRYQPLGVHDKSQLI-RTDYQILD---LGKVKIEDLIIYELHVGTFSQEGNFKGVI 123
           |||  :| || |||  |:    : |   |    |: :::  :|||||:|||: ||  :
Db      78 PDPRTRRQPDGVHALSRTFDPSAYSWQDDAWQGR-ELQGAVIYELHLGTFTPEGTLEAAA 136

Qy     124 EKLDYLKDLGITGIELMPVAQFPGNRDWGYDGVFLYAVQNTYGGPWELAKLVNEAHKRGI 183
            |||||  ||:  |||:||  | |  :||||||  :||   ||||    : |: ||  |:
Db     137 GKLDYLAGLGVDFIELLPVNAFNGTHNWGYDGVQWFAVHEAYGGPEAYQRFVDAAHAAGL 196

Qy     184 AVILDVVYNHIGPEGNYLLGLGPYFSDRYKTPWGLTFNFDDRGCDQVRKFILENVEYWFK 243
            || |||||:|| ||||   |||        || : | |  | | ||.::||:|:  | :
Db     197 GVIQDVVYNHLGPSGNYLPRFGPYLKQGEGNTWGDSVNLDGPGSDHVRRYILDNLAMWLR 256

Qy     244 TFKIDGLRLDAVHAIFDNSPKHILQEIAEKAHQLGKFV------IAESDLNDPKIV--KD 295
            :::|||||||||||: |    |||::    | |:   |      |||||||:|:::  :|
Db     257 DYRVDGLRLDAVHALKDERAVHILEDFGALADQISAEVGRPLTLIAESDLNNPRLLYPRD 316

Qy     296 DCGYKIDAQWVDDFHHAVHAFITKEKDYYYQDFGRIEDIEKTFKDVFVYDGKYSRYRGRT 355
             || :: || |||||||||  :| |   || ||  :  : |  :| | :|| || :| |
Db     317 VNGYGLEGQWSDDFHHAVHVNVTGETTGYYSDFDSLAALAKVLRDGFFHDGSYSSFRERH 376

Qy     356 HGAPV--GDLPPRKFVVFIQNHDQVGNRGNGERLSILTDKTTYLMAATLYILSPYIPLIF 413
           || |:    : |   ||  |||||:|||  |:|||      :  :|| | :  |: |::
Db     377 HGRPINFSAVHPAALVVCSQNHDQIGNRATGDRLSQTLPYGSLALAAVLTLTGPFTPMLL 436

Qy     414 MGEEYYETNPFFFFSDFSDPVLIKGVREGRLKENNQM-------IDPQSEEAFLKSKLSW 466
           |||||  : |: ||:   :| | |   |||:||  :|        |||  | | :||| |
Db     437 MGEEYGASTPWQFFTSHPEPELGKATAEGRIKEFERMGWDPAVVPDPQDPETFRRSKLDW 496

Qy     467 KIDEE-----VLDYYKQLINIRK 484
               |     :|: |: |  :|:
Db     497 AEAAEGDHARLLELYRSLTALRR 519
```

Search completed: July 15, 2002, 12:18:05
Job time: 97 sec

GenCore version 4.5

OM protein - protein search, using sw model

Run on: July 15, 2002, 12:18:43 ; Search time 32.43 Seconds
(without alignments)
1647.415 Million cell updates/sec

Title: US-09-298-924-8
Perfect score: 2954
Sequence: 1 MFSFGGNIEKNKGIFKLWAP.....KLKKDELIKVNRGVGVYQLE 556

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2954	100.0	556	2	JC5132	alpha-amylase (EC
2	1784.5	60.4	561	2	S73087	alpha-amylase (EC
3	1748.5	59.2	558	2	JC5135	alpha-amylase (EC
4	1212.5	41.0	620	2	AH1827	alpha-amylase [imp
5	969	32.8	596	2	JC4697	trehalose trehaloh
6	959	32.5	598	2	S65770	maltooligosyltreha
7	936	31.7	575	2	JC7727	malto-oligosyltreh
8	900.5	30.5	594	2	AI0673	probable hydrolase
9	866.5	29.3	580	2	G70763	probable maltoolig
10	865.5	29.3	583	2	G83375	probable glycosyl
11	835.5	28.3	600	2	H75516	probable maltoolig
12	793.5	26.9	595	2	AG3194	alpha-amylase Atu5
13	792.5	26.8	601	2	D96001	probable 1,4-alpha
14	493	16.7	552	2	S19134	probable alpha-glu
15	493	16.7	552	2	AH1915	hypothetical prote
16	353	11.9	666	2	B56639	1,4-alpha-glucan b
17	342.5	11.6	638	2	S18599	1,4-alpha-glucan b
18	333.5	11.3	707	2	S77094	glycogen operon pr
19	332.5	11.3	1280	2	E95031	alkaline amylopull
20	331.5	11.2	1256	2	G97902	alpha-amylase (EC
21	326	11.0	718	2	G69585	pullulanase amyX -
22	322.5	10.9	422	2	S31839	1,4-alpha-glucan b
23	320	10.8	843	2	H72204	pullulanase - Ther
24	319.5	10.8	718	2	S73088	glycogen operon pr
25	315	10.7	562	2	S01312	alpha-amylase (EC
26	313.5	10.6	627	1	S40048	1,4-alpha-glucan b
27	311.5	10.5	642	2	E98000	1,4-alpha-glucan b
28	310.5	10.5	642	2	G95129	1,4-alpha-glucan b
29	306.5	10.4	666	2	E71565	probable glycogen
30	303.5	10.3	664	2	B86539	glycogen hydrolase
31	303.5	10.3	664	2	B72084	glycosyl hydrolase
32	302	10.2	783	2	B84823	probable isoamylas
33	296.5	10.0	764	2	S47569	1,4-alpha-glucan b
34	296.5	10.0	818	2	T02231	probable isoamylas
35	296	10.0	630	2	D70363	1,4-alpha-glucan b
36	295	10.0	666	2	G81717	glycosyl hydrolase
37	295	10.0	789	2	T01321	probable isoamylas
38	293.5	9.9	717	2	H84057	pullulanase amyX [
39	292.5	9.9	733	2	T49679	probable branching
40	292	9.9	1072	2	A84112	alkaline amylopull
41	289	9.8	720	2	G97229	pullulanase [impor
42	289	9.8	741	1	S34218	1,4-alpha-glucan b
43	286.5	9.7	666	2	D82511	1,4-alpha-glucan b
44	283	9.6	746	2	S74546	glycogen operon pr
45	282	9.5	770	2	S76095	hypothetical prote

ALIGNMENTS

RESULT 1
JC5132
alpha-amylase (EC 3.2.1.1) - Sulfolobus acidocaldarius
N;Alternate names: glycogenase
C;Species: Sulfolobus acidocaldarius
C;Date: 31-Jan-1997 #sequence_revision 27-Feb-1997 #text_change 20-Jun-2000
C;Accession: JC5132
R;Kobayashi, K.; Kato, M.; Miura, Y.; Kettoku, M.; Komeda, T.; Iwamatsu, A.
Biosci. Biotechnol. Biochem. 60, 1720-1723, 1996
A;Title: Gene analysis of trehalose-producing enzymes from hyperthermophilic archaea
A;Reference number: JC5131; MUID:97141330
A;Accession: JC5132
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-556 <KOB>
A;Cross-references: DDBJ:D64131; NID:g987704; PIDN:BAA11011.1; PID:g987705
A;Experimental source: strain ATCC33909
C;Comment: This enzyme hydrolyzes glycosyltrehalose to liberate trehalose.
C;Superfamily: trehalose trehalohydrolase
C;Keywords: glycosidase; hydrolase

Query Match 100.0%; Score 2954; DB 2; Length 556;
Best Local Similarity 100.0%; Pred. No. 2.3e-194;
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 MFSFGGNIEKNKGIFKLWAPYVNSVKLKLSKKLIPMEKNDEGFFEVEIDDIEENLTYSYI 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MFSFGGNIEKNKGIFKLWAPYVNSVKLKLSKKLIPMEKNDEGFFEVEIDDIEENLTYSYI 60

Qy      61 IEDKREIPDPASRYQPLGVHDKSQLIRTDYQILDLGKVKIEDLIIYELHVGTFSQEGNFK 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 IEDKREIPDPASRYQPLGVHDKSQLIRTDYQILDLGKVKIEDLIIYELHVGTFSQEGNFK 120

Qy     121 GVIEKLDYLKDLGITGIELMPVAQFPGNRDWGYDGVFLYAVQNTYGGPWELAKLVNEAHK 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 GVIEKLDYLKDLGITGIELMPVAQFPGNRDWGYDGVFLYAVQNTYGGPWELAKLVNEAHK 180

Qy     181 RGIAVILDVVYNHIGPEGNYLLGLGPYFSDRYKTPWGLTFNFDDRGCDQVRKFILENVEY 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 RGIAVILDVVYNHIGPEGNYLLGLGPYFSDRYKTPWGLTFNFDDRGCDQVRKFILENVEY 240

Qy     241 WFKTFKIDGLRLDAVHAIFDNSPKHILQEIAEKAHQLGKFVIAESDLNDPKIVKDDCGYK 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 WFKTFKIDGLRLDAVHAIFDNSPKHILQEIAEKAHQLGKFVIAESDLNDPKIVKDDCGYK 300

Qy     301 IDAQWVDDFHHAVHAFITKEKDYYYQDFGRIEDIEKTFKDVFVYDGKYSRYRGRTHGAPV 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 IDAQWVDDFHHAVHAFITKEKDYYYQDFGRIEDIEKTFKDVFVYDGKYSRYRGRTHGAPV 360

Qy     361 GDLPPRKFVVFIQNHDQVGNRGNGERLSILTDKTTYLMAATLYILSPYIPLIFMGEEYYE 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    361 GDLPPRKFVVFIQNHDQVGNRGNGERLSILTDKTTYLMAATLYILSPYIPLIFMGEEYYE 420

Qy    421 TNPFFFFSDFSDPVLIKGVREGRLKENNQMIDPQSEEAFLKSKLSWKIDEEVLDYYKQLI 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 TNPFFFFSDFSDPVLIKGVREGRLKENNQMIDPQSEEAFLKSKLSWKIDEEVLDYYKQLI 480

Qy    481 NIRKRYNNCKRVKEVRREGNCITLIMEKIGIIASFDDIVINSKITGNLLIGIGFPKKLKK 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 NIRKRYNNCKRVKEVRREGNCITLIMEKIGIIASFDDIVINSKITGNLLIGIGFPKKLKK 540

Qy    541 DELIKVNRGVGVYQLE 556
          ||||||||||||||||
Db    541 DELIKVNRGVGVYQLE 556


RESULT   2
S73087
alpha-amylase (EC 3.2.1.1) precursor - Sulfolobus solfataricus
N;Alternate names: protein c0620
C;Species: Sulfolobus solfataricus
C;Date: 09-Oct-1997 #sequence_revision 24-Oct-1997 #text_change 20-Jun-2000
C;Accession: S73087
R;Sensen, C.W.; Klenk, H.P.; Singh, R.K.; Allard, G.; Chan, C.C.Y.; Liu, Q.Y.; Penny, S.
Mol. Microbiol. 22, 175-191, 1996
A;Title: Organizational characteristics and information content of an archaeal genome: 1
A;Reference number: S73076; MUID:97055432
A;Accession: S73087
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-561 <SEN>
A;Cross-references: EMBL:Y08256; NID:g1707679; PIDN:CAA69503.1; PID:g1707699
A;Experimental source: strain P2
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1996
C;Superfamily: trehalose trehalohydrolase
C;Keywords: glycosidase; hydrolase

  Query Match             60.4%;  Score 1784.5;  DB 2;  Length 561;
  Best Local Similarity   59.2%;  Pred. No. 2.4e-114;
  Matches  332;  Conservative   90;  Mismatches  130;  Indels    9;  Gaps    6;

Qy      3 SFGGNIEKNKGIFKLWAPYVNSVKLK-LSKKLIPMEKNDEGFFEVEIDDIEENLTYSYII 61
          :||  ::::   | |||||   |||| |:: :  ||::|:|:|: : :|::     | ||:
Db      2 TFGYKLDEDGVTFNLWAPYQRKVKLKILNRGIYEMERDDKGYFTITLDNVRVGDRYKYIL 61

Qy     62 EDKREIPDPASRYQPLGVHDKSQLIRTDYQILDLGKVKI--EDLIIYELHVGTFSQEGNF 119
          :|  |:|||||||||| |||  |::|  |::  |   ||:  |||:|||||:|||: || |
Db     62 DDNSEVPDPASRYQPEGVHGYSEIISPDFEWDDENSVKVKREDLVIYELHIGTFTSEGTF 121

Qy    120 KGVIEKLDYLKDLGITGIELMPVAQFPGNRDWGYDGVFLYAVQNTYGGPWELAKLVNEAH 179
          :|||:||:|||:||:| ||:||:|||| :|||||:|||||:|||     |||||||
Db    122 EGVIKKLNYLKELGVTAIEIMPIAQFPGKKDWGYDGVYLYAVQNSYGGPSGFRKLVNEAH 181

Qy    180 KRGIAVILDVVYNHIGPEGNYLLGLGPYFSDRYKTPWGLTFNFDDRGCDQVRKFILENVE 239
          | |:|||||||||:||||||:: |||||::|||||||||||||| | |:|||||||||
Db    182 KLGLAVILDVVYNHVGPEGNYMVKLGPYFSEKYKTPWGLTFNFDDAGSDEVRKFILENVE 241

Qy    240 YWFKTFKIDGLRLDAVHAIFDNSPKHILQEIAEKAHQLGKFVIAESDLNDPKIV--KDDC 297
          ||   | :|| |||||||| ||||||||::||:  |:  | ||||||||||::|  |: |
Db    242 YWINEFHVDGFRLDAVHAIIDNSPKHILEDIADVVHKYDKIVIAESDLNDPRVVNPKEKC 301

Qy    298 GYKIDAQWVDDFHHAVHAFITKEKDYYYQDFGRIEDIEKTFKDVFVYDGKYSRYRGRTHG 357
          || |||||||||||:|||:| |:  || ||| | || |::|||:|||||| :| :|||
Db    302 GYNIDAQWVDDFHHAIHAFLTGERQGYYSDFGSIGDIVKSYKDVFIYDGKYSNFRRKTHG 361

Qy    358 APVGDLPPRKFVVFIQNHDQVGNRGNGERLSILTDKTTYLMAATLYILSPYIPLIFMGEE 417
            ||||  ||||:|||||||||| |||| | || :| :|| |||||||||:||||||
Db    362 KSVGDLDGCKFVVYIQNHDQVGNRGGGERLIKLVDKESYKIAAALYILSPYIPMIFMGEE 421

Qy    418 YYETNPFFFFSDFSDPVLIKGVREGRLKENNQMIDPQSEEAFLKSKLSWKIDEEVLDYYK 477
          | | |||::|||||| ||:|||||| :|| |  ||||:  |  |||||||:::| :||
Db    422 YGEENPFYYFSDFSDPKLIQGVREGRRRENGQETDPQSDCTFNDSKLSWKINDDILSFYK 481

Qy    478 QLINIRKRYN-NCKRVKEVRREGNCITLIMEKIGIIASFDDIVINSKITGNLLI--GIGF 534
          || ||| |   | |   |       :|:       :  |   ||  | :| |::     |
Db    482 SLIKIRKEYGLACNRKLSVENGNYWLTVKGNGCLAVYVFSKSVIEMKYSGTLVLSSNSSF 541

Qy    535 PKKLKKDELIKVNRGVGVYQL 555
          | :: : :  ::::|  :|:|
Db    542 PSQITESK-YELDKGFALYKL 561


RESULT   3
JC5135
alpha-amylase (EC 3.2.1.1) - Sulfolobus solfataricus
C;Species: Sulfolobus solfataricus
C;Date: 31-Jan-1997 #sequence_revision 27-Feb-1997 #text_change 20-Jun-2000
C;Accession: JC5135
R;Kobayashi, K.; Kato, M.; Miura, Y.; Kettoku, M.; Komeda, T.; Iwamatsu, A.
Biosci. Biotechnol. Biochem. 60, 1882-1885, 1996
A;Title: Gene cloning and expression of new trehalose-producing enzymes from the hype
A;Reference number: JC5134; MUID:97141610
A;Accession: JC5135
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-558 <KOB>
A;Cross-references: DDBJ:D64130; NID:g987702; PIDN:BAA11010.1; PID:g1838936
A;Experimental source: strain KM1
C;Comment: This enzyme hydrolyzes starch alpha-1,4 linkage adjacent to the alpha-1,1
C;Superfamily: trehalose trehalohydrolase
C;Keywords: glycosidase; hydrolase

  Query Match             59.2%;  Score 1748.5;  DB 2;  Length 558;
  Best Local Similarity   58.9%;  Pred. No. 6.9e-112;
  Matches  329;  Conservative   87;  Mismatches  136;  Indels    7;  Gaps    5;

Qy      3 SFGGNIEKNKGIFKLWAPYVNSVKLK-LSKKLIPMEKNDEGFFEVEIDDIEENLTYSYII 61
          :|   |: |: || |||||  ||||| | | |  ||::::|:| : :::::    | |::
Db      1 TFAYKIDGNEVIFTLWAPYQKSVKLKVLEKGLYEMERDEKGYFTITLNNVKVRDRYKYVL 60

Qy     62 EDKREIPDPASRYQPLGVHDKSQLIRTDYQILDLGKVKIEDLIIYELHVGTFSQEGNFKG 121
          :|  ||||||||||| |||  ||:|:   :  :  :| |||||||:|||||: || |:|
Db     61 DDASEIPDPASRYQPEGVHGPSQIIQESKEFNNETFLKKEDLIIYEIHVGTFTPEGTFEG 120

Qy    122 VIEKLDYLKDLGITGIELMPVAQFPGNRDWGYDGVFLYAVQNTYGGPWELAKLVNEAHKR 181
          || |||||||||| ||:||:||||| ||||||||:||||||:||||    |||:||||:
Db    121 VIRKLDYLKDLGITAIEIMPIAQFPGKRDWGYDGVYLYAVQNSYGGPEGFRKLVDEAHKK 180

Qy    182 GIAVILDVVYNHIGPEGNYLLGLGPYFSDRYKTPWGLTFNFDDRGCDQVRKFILENVEYW 241
          |: |||||||||:|||||:: |||||  :|||||||||||||   |:||||||||||||
Db    181 GLGVILDVVYNHVGPEGNYMVKLGPYFSQKYKTPWGLTFNFDDAESDEVRKFILENVEYW 240

Qy    242 FKTFKIDGLRLDAVHAIFDNSPKHILQEIAEKAHQLGKFVIAESDLNDPKIV--KDDCGY 299
          | : :|| |||||||| | ||||||:|||:  |:  : ||||||||||::|  |: |||
Db    241 IKEYNVDGFRLDAVHAIIDTSPKHILEEIADVVHKYNRIVIAESDLNDPRVVNPKEKCGY 300

Qy    300 KIDAQWVDDFHHAVHAFITKEKDYYYQDFGRIEDIEKTFKDVFVYDGKYSRYRGRTHGAP 359
          |||||||||||::||::| |:  || ||| ::|| |::||||||||||| :| :||| |
Db    301 NIDAQWVDDFHHSIHAYLTGERQGYYTDFGNLDDIVKSYKDVFVYDGKYSNFRRKTHGEP 360

Qy    360 VGDLPPRKFVVFIQNHDQVGNRGNGERLSILTDKTTYLMAATLYILSPYIPLIFMGEEYY 419
          ||:|    |||:||||||||||| |||:  | |: :| :|| ||:||||||:|||||||
Db    361 VGELDGCNFVVYIQNHDQVGNRGKGERIIKLVDRESYKIAAALYLLSPYIPMIFMGEEYG 420

Qy    420 ETNPFFFFSDFSDPVLIKGVREGRLKENNQMIDPQSEEAFLKSKLSWKIDEEVLDYYKQL 479
          | |||:|||||||  ||:|||||| ||| |  ||| |  |  |||||||||:  :|| |
Db    421 EENPFYFFSDFSDSKLIQGVREGRKKENGQDTDPQDESTFNASKLSWKIDEEIFSFYKIL 480

Qy    480 INIRKRYN-NCKRVKEVRREGNCITLIMEKIGIIASFDDIVINSKITGNLLIGI--GFPK 536
          | :||  :  | |   |     | : :   :   :  |     |  | :| ||:     ||:
Db    481 IKMRKELSIACDRRVNVVNGENWLIIKGREYFSLYVFSKSSIEVKYSGTLLLSSNNSFPQ 540

Qy    537 KLKKDELIKVNRGVGVYQL 555
          ::: :  : ::|  :|:|
Db    541 HIEEGK-YEFDKGFALYKL 558
```

```
RESULT   4
AH1827
alpha-amylase [imported] - Anabaena sp. (strain PCC 7120)
C;Species: Anabaena sp.
A;Note: Anabaena sp. (strain PCC 7120) is a synonym of Nostoc sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 11-Jan-2002
C;Accession: AH1827
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
 Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AH1827
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-620 <KUR>
A;Cross-references: GB:BA000019; PIDN:BAB77692.1; PID:g17135146; GSPDB:GN00179
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: all0168
C;Superfamily: trehalose trehalohydrolase

  Query Match             41.0%;  Score 1212.5;  DB 2;  Length 620;
  Best Local Similarity   46.5%;  Pred. No. 3.7e-75;
  Matches  235;  Conservative   86;  Mismatches  151;  Indels   33;  Gaps    7;

Qy      15 FKLWAPYVNSVK---LKLSKKLIPMEKNDEGFFEVEIDDIEENLTYSYIIEDKREIPDPA 71
           | :|:| :|||    ||  :||||::   ||:::::::|:     | | : |:    |||
Db      29 FTVWSPTLNSVAVQILKPEEKLIPLKAQGEGYWQIKVNDVYPGTLYRYQLNDQEAFADPA 88

Qy      72 SRYQPLGVHDKSQLIRTDYQILD--LGKVKIEDLIIYELHVGTFSQEGNFKGVIEKLDYL 129
           |:||| |||  ||::   ::  |     : :| :| ||||||||: || |  :| :|  |
Db      89 SQYQPEGVHGASQVVDHKFEWTDKTWSGISLESMIFYELHVGTFTPEGTFTTIIPRLPEL 148

Qy     130 KDLGITGIELMPVAQFPGN---------RDWGYDGVFLYAVQNTYGGPWELAKLVNEAHK 180
           ::|||  |||||:|||||:         |:||||||: :||||:|| | :|   ||  |:
Db     149 RELGINAIELMPIAQFPGDDHIEPDLAYRNWGYDGVYPFAVQNSYGSPADLKNFVNACHE 208

Qy     181 RGIAVILDVVYNHIGPEGNYLLGLGPYFSDRYKTPWGLTFNFDDRGCDQVRKFILENVEY 240
           ||||:|||||||| ||||||:    |||:  |||||  ||||     || : ::|  |
Db     209 NGIAVVLDVVYNHFGPEGNYMGQFAPYFTRTYKTPWGNAMNFDDAYSQGVRNYFIQNALY 268

Qy     241 WFKTFKIDGLRLDAVHAIFDNSPKHILQEIAEKAHQLG------KFVIAESDLNDPKIVK 294
           |   | ||||||||: ||:|   || | |:||  |         : :|||||||:|:|::
Db     269 WLGEFHIDGLRLDAIQAIYDLGAKHFLWELAEAVHNFSQGGTWKRHLIAESDLNNPQIIR 328

Qy     295 --DDCGYKIDAQWVDDFHHAVHAFITKEKDYYYQDFGRIEDIEKTFKDVFVYDGKYSRYR 352
             :  || :|||| ||||||:|| :| ::  ||||||:  |: | : | |||| :|: :|
Db     329 PVESGGYGLDAQWSDDFHHALHALLTGDRQGYYQDFGKCADLAKAYADTFVYDWRYAPHR 388

Qy     353 GRTHGAPVGDLPPRKFVVFIQNHDQVGNRGNGERLSILTDKTTYLMAATLYILSPYIPLI 412
           | ||    | |  :| | |||||:||:  |||||          :||   :||||:||:
Db     389 KRFHGISCRDRPLSQFSVCIQNHDQIGNQMQGERLSERISFAGLKLAAGAVLLSPYLPLL 448

Qy     413 FMGEEYYETNPFFFFSDFSDPVLIKGVREGRLKE------NNQMIDPQSEEAFLKSKLSW 466
           |||||| || || :|   ||  ||: || || :|          ||:| | ||: ||:|
Db     449 FMGEEYGETAPFIYFVSHSDLDLIQAVRAGRKEEFEAFHYAEDPPDPESAETFLRCKLNW 508

Qy     467 KIDEE-----VLDYYKQLINIRKRY 486
           ::  :     : |:|:|||::|| :
Db     509 ELRHQGQHKVLWDWYRQLIHLRKTH 533


RESULT   5
JC4697
trehalose trehalohydrolase (EC 2.4.1.-) - Rhizobium sp.
C;Species: Rhizobium sp.
C;Date: 10-May-1996 #sequence_revision 19-Jul-1996 #text_change 20-Jun-2000
C;Accession: JC4697
R;Maruta, K.; Hattori, K.; Nakada, T.; Kubota, M.; Sugimoto, T.; Kurimoto, M.
Biosci. Biotechnol. Biochem. 60, 717-720, 1996
A;Title: Cloning and sequencing of trehalose biosynthesis genes from Rhizobium sp. M-
A;Reference number: JC4696; MUID:96219094
A;Accession: JC4697
A;Molecule type: DNA
A;Residues: 1-596 <MAR>
A;Cross-references: DDBJ:D78001; NID:g1040695; PIDN:BAA11187.1; PID:g1040697
C;Comment: This enzyme belongs to the alpha-amylase family. It hydrolyzes the produc
oligosyl group and trehalose.
C;Genetics:
A;Gene: MTHase
C;Superfamily: trehalose trehalohydrolase
C;Keywords: glycosyltransferase; hexosyltransferase

  Query Match             32.8%;  Score 969;  DB 2;  Length 596;
  Best Local Similarity   39.5%;  Pred. No. 1.6e-58;
  Matches  215;  Conservative   74;  Mismatches  195;  Indels   60;  Gaps   12;

Qy      13 GIFKLWAPYVNSVKLKLSKKLIPMEKN------DEGFFEVEIDDIEENLTYSYIIE-DKR 65
           | | :|||   :| |    :  | :       |||::         :: | |::: |:
Db      15 GRFDIWAPEAGTVTLLAGGERYEMGRRPGNGPADEGWWTAADAPTGADVDYGYLLDGDEI 74

Qy      66 EIPDPASRYQPLGVHDKSQLIRTDYQILDLGKVKIEDL----------IIYELHVGTFSQ 115
            :||| :| || |||    | ||     | |  : :|          :|||||:|||:
Db      75 PLPDPRTRRQPEGVH---ALSRT----FDPGAHRWQDAGWQGRELQGSVIYELHIGTFTP 127

Qy     116 EGNFKGVIEKLDYLKDLGITGIELMPVAQFPGNRDWGYDGVFLYAVQNTYGGPWELAKLV 175
           ||       |||||  |||  |||:||  | |  :|||||  :||   ||||    : |
Db     128 EGTLDAAAGKLDYLAGLGIDFIELLPVNAFNGTHNWGYDGVQWFAVHEGYGGPAAYQRFV 187

Qy     176 NEAHKRGIAVILDVVYNHIGPEGNYLLGLGPYFSDRYKTPWGLTFNFDDRGCDQVRKFIL 235
           : ||  |: || ||||||:|| ||||   |||         || : | |  | | ||::||
Db     188 DAAHAAGLGVIQDVVYNHLGPSGNYLPRYGPYLKHGEGNTWGDSVNLDGPGSDHVRQYIL 247

Qy     236 ENVEYWFKTFKIDGLRLDAVHAIFDNSPKHILQEIAEKAHQLGK------FVIAESDLND 289
           :||  | : :::|||||||||: |    |||:|    |  |         :|||||||:
Db     248 DNVAMWLRDYRVDGLRLDAVHALKDERAVHILEEFGALADALSSEGGRPLTLIAESDLNN 307

Qy     290 PKIV--KDDCGYKIDAQWVDDFHHAVHAFITKEKDYYYQDFGRIEDIEKTFKDVFVYDGK 347
           |:::  :|  || :  || |||||||  :: |   || ||  :  : |  :| | :||
Db     308 PRLLYPRDVNGYGLAGQWSDDFHHAVHVNVSGETTGYYSDFDSLGALAKVLRDGFFHDGS 367

Qy     348 YSRYRGRTHGAPV--GDLPPRKFVVFIQNHDQVGNRGNGERLSILTDKTTYLMAATLYIL 405
           || :||| || |:    : |   ||  |||||:|||  |:|||       :  :|| | :
Db     368 YSSFRGRCHGRPINFSAVHPAALVVCSQNHDQIGNRATGDRLSQSLPYGSLALAAVLTLT 427

Qy     406 SPYIPLIFMGEEYYETNPFFFFSDFSDPVLIKGVREGRLKENNQM-------IDPQSEEA 458
           |: |::|||||  | |: ||:   :| | |   |||::|  :|       |||  |
Db     428 GPFTPMLFMGEEYGATTPWQFFTSHPEPELGKATAEGRIREFERMGWDPAVVPDPQDPET 487

Qy     459 FLKSKLSWK-----IDEEVLDYYKQLINIRKRYNNCKRVKEVRREGNCITLIMEKIGIIA 513
           | :||| |          :|: |: || :|      :   |: | |   | :
Db     488 FTRSKLDWAEASAGDHARLLELYRSLITLR------RSTPELARLGFADTAV-------- 533

Qy     514 SFDD 517
           |||
Db     534 EFDD 537


RESULT   6
S65770
maltooligosyltrehalose trehalohydrolase - Arthrobacter sp. (strain Q36)
C;Species: Arthrobacter sp.
A;Variety: strain Q36
C;Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 20-Jun-2000
C;Accession: S65770
R;Maruta, K.; Hattori, K.; Nakada, T.; Kubota, M.; Sugimoto, T.; Kurimoto, M.
Biochim. Biophys. Acta 1289, 10-13, 1996
A;Title: Cloning and sequencing of trehalose biosynthesis genes from Arthrobacter sp
A;Reference number: S65769; MUID:96195835
A;Accession: S65770
A;Status: preliminary
```

A;Molecule type: DNA
A;Residues: 1-598 <MAR>
A;Cross-references: EMBL:D63343; NID:g1255444; PIDN:BAA09668.1; PID:g1255446
C;Superfamily: trehalose trehalohydrolase

Query Match 32.5%; Score 959; DB 2; Length 598;
Best Local Similarity 40.2%; Pred. No. 7.8e-58;
Matches 202; Conservative 81; Mismatches 186; Indels 34; Gaps 10;

```
Qy      15 FKLWAPYVNSVKLKLSKKLIPMEK------NDEGFFEVEIDDIEENLTYSYIIE-DKREI 67
           : :|||   || |    :   |::       | |::       : |: | |::: |:  :
Db      19 YDVWAPNAESVTLLAGGERYAMQRRAETGPEDAGWWTAAGAPTDGNVDYGYLLDGDETPL 78

Qy      68 PDPASRYQPLGVHDKSQLI-RTDYQILD---LGKVKIEDLIIYELHVGTFSQEGNFKGVI 123
           ||| :| || |||  |:    : |   |    |: :::  :|||||:|||: ||  :
Db      79 PDPRTRRQPDGVHALSRTFDPSAYSWQDDAWQGR-ELQGAVIYELHLGTFTPEGTLEAAA 137

Qy     124 EKLDYLKDLGITGIELMPVAQFPGNRDWGYDGVFLYAVQNTYGGPWELAKLVNEAHKRGI 183
           |||||  ||:  |||:||  | |  :||||||  :||   ||||    : |: ||  |:
Db     138 GKLDYLAGLGVDFIELLPVNAFNGTHNWGYDGVQWFAVHEAYGGPEAYQRFVDAAHAAGL 197

Qy     184 AVILDVVYNHIGPEGNYLLGLGPYFSDRYKTPWGLTFNFDDRGCDQVRKFILENVEYWFK 243
           || ||||||:|| ||||   |||        || : | |  | | ||::||:|:  | :
Db     198 GVIQDVVYNHLGPSGNYLPRFGPYLKQGEGNTWGDSVNLDGPGSDHVRRYILDNLAMWLR 257

Qy     244 TFKIDGLRLDAVHAIFDNSPKHILQEIAEKAHQLGKFV------IAESDLNDPKIV--KD 295
           :::|||||||||: |    |||::    | |:   |      |||||||:|:::  :|
Db     258 DYRVDGLRLDAVHALKDERAVHILEDFGALADQISAEVGRPLTLIAESDLNNPRLLYPRD 317

Qy     296 DCGYKIDAQWVDDFHHAVHAFITKEKDYYYQDFGRIEDIEKTFKDVFVYDGKYSRYRGRT 355
           || :: || |||||||| :| |   || ||  :  : |  :| | :|| || :| |
Db     318 VNGYGLEGQWSDDFHHAVHVNVTGETTGYYSDFDSLAALAKVLRDGFFHDGSYSSFRERH 377

Qy     356 HGAPV--GDLPPRKFVVFIQNHDQVGNRGNGERLSILTDKTTYLMAATLYILSPYIPLIF 413
           || |:    : |   ||  |||||:|||  |:|||      :  :|| | :  |: |::
Db     378 HGRPINFSAVHPAALVVCSQNHDQIGNRATGDRLSQTLPYGSLALAAVLTLTGPFTPMLL 437

Qy     414 MGEEYYETNPFFFFSDFSDPVLIKGVREGRLKENNQM-------IDPQSEEAFLKSKLSW 466
           |||||  : |: ||:   :| | |   |||:||  :|        |||  | | :||| |
Db     438 MGEEYGASTPWQFFTSHPEPELGKATAEGRIKEFERMGWDPAVVPDPQDPETFRRSKLDW 497

Qy     467 KIDEE-----VLDYYKQLINIRK 484
               |     :|: |: |  :|:
Db     498 AEAAEGDHARLLELYRSLTALRR 520
```

RESULT 7
JC7727
malto-oligosyltrehalose trehalohydrolase (EC 3.2.1.141) - Arthrobacter ramosus
C;Species: Arthrobacter ramosus
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 31-Dec-2001
C;Accession: JC7727; PC7175
R;Yamamoto, T.; Maruta, K.; Watanabe, H.; Yamashita, H.; Kubota, M.; Fukuda, S.; Kurimot
Biosci. Biotechnol. Biochem. 65, 1419-1423, 2001
A;Title: Trehalose-producing operon treYZ from Arthrobacter ramosus S34.
A;Reference number: JC7726; MUID:21364352; PMID:11471747
A;Accession: JC7727
A;Molecule type: DNA
A;Residues: 1-575 <YAM>
A;Cross-references: DDBJ:AB045141
A;Experimental source: strain S34
A;Accession: PC7175
A;Molecule type: protein
A;Residues: 1-20;31-50;298-319 <YA2>
C;Comment: This enzyme is involved in producing trehalose from glycogen.
C;Genetics:
A;Gene: treZ
C;Superfamily: trehalose trehalohydrolase
C;Keywords: glycosidase; hydrolase

Query Match 31.7%; Score 936; DB 2; Length 575;
Best Local Similarity 38.7%; Pred. No. 2.8e-56;
Matches 198; Conservative 82; Mismatches 177; Indels 54; Gaps 10;

```
Qy      15 FKLWAPYVNSVKLKL--SKKLIPMEKNDEGFFEVE--IDDIEENLTYSYIIEDKREIPDP 70
           | :|||    | | :   :  :|: ::: |:: ::   |   : : | |::: |    ||
Db       5 FPVWAPQAAQVTLVVGQGRAELPLTRDENGWWALQQPWDGGPDLVDYGYLVDGKGPFADP 64

Qy      71 ASRYQPLGVHDKSQLIRTDYQILDLGKV----------------KIEDLIIYELHVGTFS 114
           |  || |||              :||:                    :   :|||||||||:
Db      65 RSLRQPRGVH-------------ELGREFDPARYAWGDDGWRGRDLTGAVIYELHVGTFT 111

Qy     115 QEGNFKGVIEKLDYLKDLGITGIELMPVAQFPGNRDWGYDGVFLYAVQNTYGGPWELAKL 174
           ||     | :||:|  ||:  :||:||  | |   ||||||  |||   ||||    :
Db     112 PEGTLDSAIRRLDHLVRLGVDAVELLPVNAFNGTHGWGYDGVLWYAVHEPYGGPEAYQRF 171

Qy     175 VNEAHKRGIAVILDVVYNHIGPEGNYLLGLGPYFSDRYKTPWGLTFNFDDRGCDQVRKFI 234
           |:  | ||:||: ||||||:|| ||:|   |||         ||   | |    |:||::|
Db     172 VDACHARGLAVVQDVVYNHLGPSGNHLPDFGPYLGSGAANTWGDALNLDGPLSDEVRRYI 231

Qy     235 LENVEYWFKTFKIDGLRLDAVHAIFDNSPKHILQEIA----EKAHQLGK--FVIAESDLN 288
           ::|  || :    ||||||||||: |    |:|:|:|    | | :||:   :|||||||
Db     232 IDNAVYWLRDMHADGLRLDAVHALRDARALHLLEELAARVDELAGELGRPLTLIAESDLN 291

Qy     289 DPKIVKDDC--GYKIDAQWVDDFHHAVHAFITKEKDYYYQDFGRIEDIEKTFKDVFVYDG 346
           |||:::     || :|||| || |||||| :| |   || ||| :  : | |:  : :||
Db     292 DPKLIRSRAAHGYGLDAQWDDDVHHAVHANVTGETVGYYADFGGLGALVKVFQRGWFHDG 351

Qy     347 KYSRYRGRTHGAPVG-DLPPRKFVVFIQNHDQVGNRGNGERLSILTDKTTYLMAATLYIL 405
           :| :| | || |:  |:| |: | | |:|||||| |:|:|    : :   || | :|
Db     352 TWSSFRERHHGRPLDPDIPFRRLVAFAQDHDQVGNRAVGDRMSAQVGEGSLAAAAALVLL 411

Qy     406 SPYIPLIFMGEEYYETNPFFFFSDFSDPVLIKGVREGRLKENNQM-------IDPQSEEA 458
           |: |::|||||:    |: ||:   :| | :    ||: |  :|        |||
Db     412 GPFTPMLFMGEEWGARTPWQFFTSHPEPELGEATARGRIAEFARMGWDPAVVPDPQDPAT 471

Qy     459 FLKSKLSWKIDEE-----VLDYYKQLINIRK 484
           | :| | |   |       :| :|  || :|:
Db     472 FARSHLDWSEPEREPHAGLLAFYTDLIALRR 502
```

RESULT 8
AI0673
probable hydrolase STY1503 [imported] - Salmonella enterica subsp. enterica serovar T
C;Species: Salmonella enterica subsp. enterica serovar Typhi
A;Note: this species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 27-Nov-2001
C;Accession: AI0673
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Church
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farr
, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens,
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica se
A;Reference number: AB0502; PMID:11677608
A;Accession: AI0673
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-594 <PAR>
A;Cross-references: GB:AL513382; PIDN:CAD01762.1; PID:g16502610; GSPDB:GN00176
C;Genetics:
A;Gene: STY1503
C;Superfamily: trehalose trehalohydrolase

Query Match 30.5%; Score 900.5; DB 2; Length 594;
Best Local Similarity 38.9%; Pred. No. 7.8e-54;
Matches 199; Conservative 79; Mismatches 173; Indels 61; Gaps 13;

```
Qy      15 FKLWAPYVNSVKLKLSKKLIPMEKNDEGFFEVEIDDIEENLTYSYIIEDKREIPDPASRY 74
           |:|||    | |:|: |   |: : :|:| :::  :    |:::: |   :||||||
Db      22 FRLWATGQQKVMLRLAGKDQEMQASGDGWFTLDVSGVTPGTEYNFVLSDGMVLPDPASRA 81

Qy      75 QPLGVHDKSQLIRTDYQILDLGKV----------KIEDLIIYELHVGTFSQEGNFKGVIE 124
```

```
Db     82 QKTDVNGPS-------YVIDPGSYAWRNTGWKGSRWEQAVVYEMHTGTFTPEGTFHTAIA 134

Qy    125 KLDYLKDLGITGIELMPVAQFPGNRDWGYDGVFLYAVQNTYGGPWELAKLVNEAHKRGIA 184
Db    135 KLPYLAELGVTVIEVMPVAQFGGERGWGYDGVLLYAPHSAYGTPDDFKAFIDTAHGYGLS 194

Qy    185 VILDVVYNHIGPEGNYLLGLGP-YFSDRYKTPWGLTFNFDDRGCDQVRKFILENVEYWFK 243
Db    195 VILDIVLNHFGPEGNYLPLLAPAFFHKERMTPWGNGIAYD---VDAVRRYIIEAPLYWLT 251

Qy    244 TFKIDGLRLDAVHAIFDNSPKHILQEIAEKAHQLGKFVIAESDLNDPKI---------- 292
Db    252 EYHLDGLRFDAIDQIEDSSARHVLVEIAQRIRE---------DITDRPIHLTTEDSRNII 302

Qy    293 ---VKDDCGYK--IDAQWVDDFHHAVHAFITKEKDYYYQDFGRIED--IEKTFKDVFVYD 345
Db    303 SLHPRDQDGNAPLFTAEWNDDFHNAVHVFATGETQAYYNDFADAPEKHLARALAEGFAYQ 362

Qy    346 GKYSRYRGRTHGAPVGDLPPRKFVVFIQNHDQVGNRGNGERLSILTD-KTTYLMAATLYI 404
Db    363 GEISPQTGEPRGVKSTGQPPVAFVDFIQNHDQVGNRAQGDRLITLAGAERTKVLLATL-L 421

Qy    405 LSPYIPLIFMGEEYYETNPFFFFSDFSDPVLIKGVREGRLKE-----NNQMIDPQSEEAF 459
Db    422 LSPHIPLLFMGEEYGESRPFLFFTDFHGD-LARAVREGRAKEFADHAGENVPDPNAPETF 480

Qy    460 LKSKLSWKIDEEV-----LDYYKQLINIRKRY 486
Db    481 QRSKLNWKQQHSEEGKAWLAFTRELLLLRQKH 512
```

RESULT 9
G70763
probable maltooligosyltrehalose trehalohydrolase - Mycobacterium tuberculosis (strain H3
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C;Accession: G70763
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Sqares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: A70500; MUID:98295987
A;Accession: G70763
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-580 <COL>
A;Cross-references: GB:Z74020; GB:AL123456; NID:g3261584; PIDN:CAA98329.1; PID:g1403480
A;Experimental source: strain H37Rv
C;Genetics:
A;Gene: glgZ
C;Superfamily: trehalose trehalohydrolase

  Query Match   29.3%; Score 866.5; DB 2; Length 580;
  Best Local Similarity   37.3%; Pred. No. 1.6e-51;
  Matches 187; Conservative   83; Mismatches 194; Indels   37; Gaps   11;

```
Qy     15 FKLWAPYVNSVKLKLSKKLIPMEKNDEGFFEVEIDDIEENLTYSYIIEDKREI-PDPASR 73
Db      4 FRVWAPKPALVRLDVNGAVHAMTRSADGWWHTTV-AAPADARYGYLLDDDPTVLPDPRSA 62

Qy     74 YQPLGVHDKSQ---------LIRTDYQILDLGKVKIEDLIIYELHVGTFSQEGNFKGVIE 124
Db     63 RQPDGVHARSQRWEPPGQFGAARTD---TGWPGRSVEGAVIYELHIGTFTTAGTFDAAIE 119

Qy    125 KLDYLKDLGITGIELMPVAQFPGNRDWGYDGVFLYAVQNTYGGPWELAKLVNEAHKRGIA 184
Db    120 KLDYLVDLGIDFVELMPVNSFAGTRGWGYDGVLWYSVHEPYGGPDGLVRFIDACHARRLG 179

Qy    185 VILDVVYNHIGPEGNYLLGLGPYFSDRYKTPWGLTFNFDDRGCDQVRKFILENVEYWFKT 244
Db    180 VLIDAVFNHLGPSGNYLPRFGPYLSSA-SNPWGDGINIAGADSDEVRHYIIDCALRWMRD 238

Qy    245 FKIDGLRLDAVHAIFDNSPKHILQEIAEKAH----QLGK--FVIAESDLNDPKIV--KDD 296
Db    239 FHADGLRLDAVHALVDTTAVHVLEELANATRWLSGQLGRPLSLIAETDRNDPRLITRPSH 298

Qy    297 CGYKIDAQWVDDFHHAVHAFITKEKDYYYQDFGRIEDIEKTFKDVFVYDGKYSRYRGRTH 356
Db    299 GGYGITAQWNDDIHHAIHTAVSGERQGYYADFGSLATLAYTLRNGYFHAGTYSSFRRRRH 358

Qy    357 GAPV--GDLPPRKFVVFIQNHDQVGNRGNGERLSILTDKTTYLMAATLYILSPYIPLIFM 414
Db    359 GRALDTSAIPATRLLAYTCTHDQVGNRALGDRPSQYLTGGQLAIKAALTLGSPYTAMLFM 418

Qy    415 GEEYYETNPFFFFSDFSDPVLIKGVREGRLKE-------NNQMIDPQSEEAFLKSKLSWK 467
Db    419 GEEWGASSPFQFFCSHPEPELAHSTVAGRKEEFAEHGWAADDIPDPQDPQTFQRCKLNWA 478

Qy    468 -----IDEEVLDYYKQLINIR 483
Db    479 EAGSGEHARLHRFYRDLIALR 499
```

RESULT 10
G83375
probable glycosyl hydrolase PA2164 [imported] - Pseudomonas aeruginosa (strain PAO1)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C;Accession: G83375
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, M.J.;
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa
A;Reference number: A82950; MUID:20437337
A;Accession: G83375
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-583 <STO>
A;Cross-references: GB:AE004643; GB:AE004091; NID:g9948178; PIDN:AAG05552.1; GSPDB:GN
A;Experimental source: strain PAO1
C;Genetics:
A;Gene: PA2164
C;Superfamily: trehalose trehalohydrolase

  Query Match   29.3%; Score 865.5; DB 2; Length 583;
  Best Local Similarity   38.2%; Pred. No. 1.9e-51;
  Matches 199; Conservative   89; Mismatches 166; Indels   67; Gaps   20;

```
Qy      4 FGGNIEKN-KGIFKLWAPYVNSVKLKLSK-KLIPMEKNDEGFFEVEIDDIEENLTYSYII 61
Db      7 FGAQFQGNGRTCFGLWAPDAREVRVETADGRDWPLEGSDSGWFEATL-PCPPGTRYRYRI 65

Qy     62 EDKREIPDPASRYQPLGVHDKSQLIRTDYQILDLGKV--KIE--------DLIIYELHVG 111
Db     66 DGRPGVPDPASQFQPDGVHGHS-------QVLDHGTYAWRVDEWRGRPWHEAVIYELHVG 118

Qy    112 TFSQEGNFKGVIEKLDYLKDLGITGIELMPVAQFPGNRDWGYDGVFLYAVQNTYGGPWEL 171
Db    119 LF---GSYAEVERFLPRLVELGVTAVELMPLGEFPGRRNWGYDGVLPFAPASAYGTPEQL 175

Qy    172 AKLVNEAHKRGIAVILDVVYNHIGPEGNYLLG-LGPYFSDRYKTPWGLTFNFDDRGCDQV 230
Db    176 KHLIDSAHGMGLMVFVDVIYNHFGPDGNYLAQYAAAFFRDDRQTPWGQAIDF-RRG--EV 232

Qy    231 RKFILENVEYWFKTFKIDGLRLDAVHAIFDNSPKHILQEIAEKAH-----QLGKFVIAES 285
Db    233 REFFYENALMWLLDYRVDGLRFDAVHAIPDSA---FLVEMARRLRGAAGPERHVHLVLEN 289

Qy    286 DLNDPKIVKDDCGYKIDAQWVDDFHHAVHAFITKEKDYYYQDFGR-IEDIEKTFKDVFVY 344
```

```
Db    290 DDNRASLLRQ--GY--DAQWNDDGHHALHVLLTGENDGYYQDYPEPLRCLARCLAEGFVY 345

Qy    345 DGKYSRYRGRTHGAPVGDLPPRKFVVFIQNHDQVGNRGNGERLSILTDKTTYLMAATLYI 404
          |: :|: || | |  || |  ||:|:|||||||||  |||||:| :     :|  | :
Db    346 QGEANRH-GRPRGEPSADLAPDAFVLFLQNHDQVGNRAFGERLSVLAEPQALRLAIALQL 404

Qy    405 LSPYIPLIFMGEEYYETNPFFFFSDFSDPVLIKGVREGRLKENNQ-----------MID 452
          |:| |||:|||||     || :|:|     |   ||||| ||  :           : |
Db    405 LAPMIPLLFMGEECAAREPFLYFTDHQGE-LADAVREGRRKEFGEFGRFGEGATLASLPD 463

Qy    453 PQSEEAFLKSK-------LSWKIDEEVLDYYKQLINIRKRY 486
          | : | | :|:        :|:       :|:||: ||  :
Db    464 PNAVETFERSRPGLADCDPAWR------GFYRQLLEIRHEH 498


RESULT  11
H75516
probable maltooligosyltrehalose trehalohydrolase - Deinococcus radiodurans (strain R1)
C;Species: Deinococcus radiodurans
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
C;Accession: H75516
R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
, M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A;Reference number: A75250; MUID:20036896
A;Accession: H75516
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-600 <WHI>
A;Cross-references: GB:AE001905; GB:AE000513; NID:g6458144; PIDN:AAF10042.1; PID:g645814
A;Experimental source: strain R1
C;Genetics:
A;Gene: DR0464
A;Map position: 1
C;Superfamily: trehalose trehalohydrolase


  Query Match             28.3%;  Score 835.5;  DB 2;  Length 600;
  Best Local Similarity   36.8%;  Pred. No. 2.2e-49;
  Matches  181;  Conservative   89;  Mismatches  191;  Indels   31;  Gaps   12;

Qy     15 FKLWAPYVNSVKLKLSKKLIPMEKNDEGFFEVEIDDIEENLTYSYIIEDKREIPDPASRY 74
          |:||     :| ::::     |     | :|:|:  :     | ::: |    ||| :|:
Db     38 FRLWTSTARTVAVRVNGTEHVMTSLGGGIYELEL-PVGPGARYLFVL-DGVPTPDPYARF 95

Qy     75 QPLGVHDKSQLI---RTDYQILDLGKVKIEDLIIYELHVGTFSQEGNFKGVIEKLDYLKD 131
          | ||| :::::     |:   |   :|: | : ||:|||||: || ::   ||| |||:
Db     96 LPDGVHGEAEVVDFGTFDWTDADWHGIKLADCVFYEVHVGTFTPEGTYRAAAEKLPYLKE 155

Qy    132 LGITGIELMPVAQFPGNRDWGYDGVFLYAVQNTYGGPWELAKLVNEAHKRGIAVILDVVY 191
          ||:| |::||:| | | | |||||   ||    || | :|  ||: ||: |: | |||||
Db    156 LGVTAIQVMPLAAFDGQRGWGYDGAAFYAPYAPYGRPEDLMALVDAAHRLGLGVFLDVVY 215

Qy    192 NHIGPEGNYLLGLGP-YFSDRYKTPWGLTFNFDDRGCDQVRKFILENVEYWFKTFKIDGL 250
          || || ||||    | ||:||: : ||:  :: :     :|:::  |   | : :  |||
Db    216 NHFGPSGNYLSSYAPSYFTDRFSSAWGMGLDYAE---PHMRRYVTGNARMWLRDYHFDGL 272

Qy    251 RLDAVHAIFDNSPKHILQEIAEKAHQLG--KFVIAESDLNDPKIVKDDCGYKIDAQWVDD 308
          ||||   : |:|  ||| |:|:: |:||    ::||   | | :|  :    :|  | ||
Db    273 RLDATPYMTDDSETHILTELAQEIHELGGTHLLLAEDHRNLPDLVTVN---HLDGIWTDD 329

Qy    309 FHHAVHAFITKEKDYYYQDF-GRIEDIEKTFKDVFVYDGKYSRYRGRTH--GAPVGDLPP 365
          |||     :| |:: ||  : |  | :  | :  : |:|::   :|  |  | |   |
Db    330 FHHETRVTLTGEQEGYYAGYRGGAEALAYTIRRGWRYEGQFWAVKGEEHERGHPSDALEA 389

Qy    366 RKFVVFIQNHDQVGNRGNGERLSILTDKTTYLMAATLYILSPYIPLIFMGEEYYETNPFF 425
            ||  |||||:|||  ||||      | :       :| |  ||:| |:|:  : ||
Db    390 PNFVYCIQNHDQIGNRPLGERLHQSDGVTLHEYRGAAALLLPMTPLLFQGQEWAASTPFQ 449

Qy    426 FFSDFSDPVLIKGVREGRLKE--------NNQMIDPQSEEAFLKSKLSWKIDE-----EV 472
          |||| .:   | : | ||| ||          : |||:|: || |||:|   |
Db    450 FFSDHAGE-LGQAVSEGRKKEFGGFSGFSGEDVPDPQAEQTFLNSKLNWAEREGGEHART 508

Qy    473 LDYYKQLINIRK 484
          |  |: |: :|:
Db    509 LRLYRDLLRLRR 520


RESULT  12
AG3194
alpha-amylase Atu5284 [imported] - Agrobacterium tumefaciens (strain C58, Dupont) pla
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 01-Feb-2002
C;Accession: AG3194
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo
erage, G.; Gillet, W.; Grant, C.; Guenthner, D.; Kutyavin, T.; Levy, R.; Li, M.; McCl
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kam
ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; PMID:11743193
A;Accession: AG3194
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-595 <KUR>
A;Cross-references: GB:AE008687; PIDN:AAL45973.1; PID:g17743726; GSPDB:GN00188
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: Atu5284
A;Genome: plasmid
C;Superfamily: trehalose trehalohydrolase


  Query Match             26.9%;  Score 793.5;  DB 2;  Length 595;
  Best Local Similarity   36.7%;  Pred. No. 1.6e-46;
  Matches  175;  Conservative   76;  Mismatches  195;  Indels   31;  Gaps   11;

Qy     15 FKLWAPYVNSVKLKLSKKLIPMEKNDEGFFEVEIDDIEENLTYSYIIEDKREIPDPASRY 74
          | :|||   :||| |:     |    :|: ::    :  :  | ::: |:  : ||||
Db     22 FSIWAPASATVKLWLNDAEFDMHTAGDGWHDITKPALPGD-RYGFVLADRTRVADPASNR 80

Qy     75 QPLGVHDKSQLIRTDYQILD---LGKVKIEDLIIYELHVGTFSQEGNFKGVIEKLDYLKD 131
          |  |    | ::  |:   :    |:    : ::||||:|||: || |    |||:|| |
Db     81 QQEGPRGPSLIVNHDFAWKNPNWKGR-PWHEAVVYELHIGTFTPEGTFAAAAEKLEYLAD 139

Qy    132 LGITGIELMPVAQFPGNRDWGYDGVFLYAVQNTYGGPWELAKLVNEAHKRGIAVILDVVY 191
          :||| |||||:| | |:| ||||||  :: |  || | ||   :::||  || |:|||||
Db    140 VGITTIELMPLATFAGSRGWGYDGVLQFSPQRDYGTPDELKAFIDQAHGHGIMVLLDVVY 199

Qy    192 NHIGPEGNYLLGLGPYFSDRYKTPWGLTFNFDDRGCDQVRKFILENVEYWFKTFKIDGLR 251
          || || || |    | |  :::|||   :|:     :|| | |:|  || :|:: ||||
Db    200 NHFGPAGNTLQAYAPAFFKKHETPWGPAPDFNRA---EVRSFFLQNAFYWLETYRFDGLR 256

Qy    252 LDAV-HAIFDNSPKHILQEIAEKAHQLGK----FVIAESDLN--DPKIVKDDCGYKIDAQ 304
          :|| |    :     | |:| :  :     :: |   |   |     |    :|||
Db    257 IDAADHLAGGDGEVDFLIEMAREVKRTIRNRHVHLVIEDARNAASPMTPMADGAILVDAQ 316

Qy    305 WVDDFHHAVHAFITKEKDYYYQDFGR--IEDIEKTFKDVFVYDGKYSRYRG-RTHGAPVG 361
          | ||||| :|   | |:   |:||     |:: ::    ||| |:    |     | | |
Db    317 WNDDFHHVIHVATTNEEGGIYEDFASRPYENLRRSLATGFVYQGEPRPSRNFAASGEPSG 376

Qy    362 DLPPRKFVVFIQNHDQVGNRGNGERLSILTDKTTYLMAATLYILSPYIPLIFMGEEYYET 421
          ||| :|| |: |||| ||| |||| |     :     : :| |  ||:|||:|:
Db    377 HLPPHRFVNFLHNHDQAGNRLRGERLRALIPPPLFGTLEAILLLCPQTPLVFMGDEHGSA 436

Qy    422 NPFFFFSDFSDPVLIKGVREGRLKE-----------NNQMIDPQSEEAFLKSKLSW 466
          |||||||| |    : :|  |||:             :  ::||  :     | | |
Db    437 NPFFFFSDHPDHNREQEIR-NRLKQAESFQGELPPDASQMVMDPNDQHTMQLSTLKW 492


RESULT  13
D96001
probable 1,4-alpha-glucan branching enzyme (EC 2.4.1.18) [imported] - Sinorhizobium m
```

```
C;Species: Sinorhizobium meliloti
C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 14-Sep-2001
C;Accession: D96001
R;Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A;Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing endo
A;Reference number: A95842; MUID:21396508; PMID:11481431
A;Accession: D96001
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-601 <KUR>
A;Cross-references: GB:AL591985; PIDN:CAC49676.1; PID:g15141163; GSPDB:GN00167
A;Experimental source: strain 1021, megaplasmid pSymB
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
 L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
hebault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A;Reference number: A96039; MUID:21368234; PMID:11474104
A;Contents: annotation
C;Genetics:
A;Gene: glgB2; SMb21447
A;Genome: plasmid
C;Superfamily: trehalose trehalohydrolase
C;Keywords: glycosyltransferase; hexosyltransferase

  Query Match             26.8%;  Score 792.5;  DB 2;  Length 601;
  Best Local Similarity   36.9%;  Pred. No. 1.9e-46;
  Matches  192;  Conservative   79;  Mismatches  199;  Indels   51;  Gaps   15;

Qy       3 SFGGN-IEKNKGIFKLWAPYVNSVKLKLSKKLIPMEKNDEGFFEVEIDDIEENLTYSYII 61
Db      13 SWGANFIDSDTCRFRLWAPDEREVDLVLGGAVHKMQSLDGGWFEITL-AAKAGERYCFRL 71

Qy      62 EDKREIPDPASRYQPLGVHDKSQLI---RTDYQILDLGKVKIEDLIIYELHVGTFSQEGN 118
Db      72 ADGTEVADPASSAQEREASGTSIVVDHAAYEWQTSSWRGRPWEEAVISELHVGCFTPEGT 131

Qy     119 FKGVIEKLDYLKDLGITGIELMPVAQFPGNRDWGYDGVFLYAVQNTYGGPWELAKLVNEA 178
Db     132 FRAAIERLPHLAGAGITAIEIMPVAQFPGVRGWGYDGVLHYAPHNAYGKPDDLKALVDAA 191

Qy     179 HKRGIAVILDVVYNHIGPEGNYLLGLGP-YFSDRYKTPWGLTFNFDDRGCDQVRKFILEN 237
Db     192 HSLGLTVLLDVVYNHFGPEGNYLSRYASRFFNKDRPTPWGASIAFEE---EAVRRYFIEN 248

Qy     238 VEYWFKTFKIDGLRLDAVHAIFDNSPKHIL----QEIAEK-AHQLGKFVIAESDLNDPKI 292
Db     249 ALYWLGHFRFDGLRLDATEQIRDTTKPHFLVALEHEVREAFAERQIHLVLEDAHRRRSLL 308

Qy     293 VKDDCGYKI--DAQWVDDFHHAVHAFITKEKDYYYQDF-----GRIEDIEKTFKDVFVYD 345
Db     309 QRDASGARMLFDAAWNDDLHNALHVVATGETKGHYRLFADEPWGK---IRSALAEGFAVP 365

Qy     346 GKYSRYRGRTHGAPVGD---LPPRKFVVFIQNHDQVGNRGNGERLSILTDKTTYLMAATL 402
Db     366 AKEDNF------SPEGSRARVPPQGRVNFLQNHDQIGNRAFGERLASLLQEDSLRVLAAM 419

Qy     403 YILSPYIPLIFMGEEYYETNPFFFFSDFSDPVLIKGVREGRL-----------KENNQM 450
Db     420 HMLTPQIPLLFMGEEYGETQPFYFFSDYQGEI-AAAIRLGRRDEAENFGGLPEGKTMDDL 478

Qy     451 IDPQSEEAFLKSKLSWK-----IDEEVLDYYKQLINIRKRY 486
Db     479 PDPLDPDVFAGSKLRWNRATSPAGERHLAYVRDLAVIRQRH 519

RESULT  14
S19134
probable alpha-glucanotransferase - Anabaena variabilis
C;Species: Anabaena variabilis
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 17-Mar-2000
C;Accession: S19134
R;Sato, N.
Plant Mol. Biol. 18, 165-170, 1992
A;Title: Cloning of a low-temperature-induced gene lti2 from the cyanobacterium Anaba
A;Reference number: S19133; MUID:92119230
A;Accession: S19134
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-552 <SAT>
A;Cross-references: EMBL:X59279; NID:g39252; PIDN:CAA41970.1; PID:g39253
C;Superfamily: neopullulanase; alpha-amylase core homology

  Query Match             16.7%;  Score 493;  DB 2;  Length 552;
  Best Local Similarity   28.6%;  Pred. No. 5.3e-26;
  Matches  159;  Conservative   76;  Mismatches  142;  Indels  178;  Gaps   26;

Qy      15 FKLWAPYVNSVKLKLS---KKLIPMEKNDEGFFE--VEIDDIEENL-----TYSYIIEDK 64
Db       7 FTLFAPYNKGAALIASFSDWQEIPMKKGDDGYFRTTVELEDGTYQYKFRVQTRSWFFEED 66

Qy      65 R--EIPDPASR--YQPLGVHDKSQLIRTDYQILDLGKVKIED--------LIIYELHVGT 112
Db      67 QWVDVTDPYATDIDESSGKDNSIARIKDGEKIVDTYVWQHDDKPLPADHELVIYELHVGD 126

Qy     113 FS-------QEGNFKGVIEKLDYLKDLGITGIELMPVAQFPGNRDWGYDGVFLYAVQNTY 165
Db     127 FSGGEDDPYARGKYKHVIEKLDYLCELGINAIELLPVKEYPGDYSWGYNPRYFFATESSY 186

Qy     166 GGPWELAKLVNEAHKRGIAVILDVVYNHI--------------------GPEGNYLLGLG 205
Db     187 GSTADLKKLVDECHQRGIRIIMDGIYNHSEASSPLTQIDHDYWYHHEPRDPDNN------ 240

Qy     206 PYFSDRYKTPWGLTFNFD--DRGCD--QVRKFILENVEYWFKTFKIDGLRLDAVHAIFDN 261
Db     241 ----------WGPEFNYEHYDENLETYPARKFIGDTVRYWVGEYHLDGIRYDAARQIANY 290

Qy     262 SPKH-ILQEIAEKAHQLGKFVIAE--------SDLNDPKIVKDDCGYKIDAQWVDDFHHA 312
Db     291 DFMHWIAQEAKKTAGAKPFYNVAEHIPETTSITNLDGP----------MDGCWHDSFYHT 340

Qy     313 VHAFITKEKDYYYQDFGRIEDIEKTFKDVFVYDGKYSRYRGRTHGAPVGDLPPRKFVVFI 372
Db     341 IKAHIC----------GDTFDLE-NLKDVI--DPKRQGFLGATN----------VVNYL 376

Qy     373 QNHD------QVGNRGNGERLSILTDK--------TTYLMAATLYILSPYIPLIFMGEEY 418
Db     377 TNHDHDHIMVELGNR------EIFHDEAFRRAKLGTAILMTAV------GVPLIWMGEEF 424

Qy     419 YETNPFFFFSDFSDPVLIKGVREGRLKENNQMIDPQSEEAFLKSKLSW-----KIDEEVL 473
Db     425 GEYKP--------------------KQQDQ-----------SKIDWTLLGNDLNRSLF 451

Qy     474 DYYKQLINIRKRYNN 488
Db     452 DYHKGLIGLRK--NN 464

RESULT  15
AH1915
hypothetical protein all0875 [imported] - Anabaena sp. (strain PCC 7120)
C;Species: Anabaena sp.
A;Note: Anabaena sp. (strain PCC 7120) is a synonym of Nostoc sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 11-Jan-2002
C;Accession: AH1915
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Irigu
 Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AH1915
```

```
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-552 <KUR>
A;Cross-references: GB:BA000019; PIDN:BAB72832.1; PID:g17130220; GSPDB:GN00179
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: all0875
C;Superfamily: neopullulanase; alpha-amylase core homology


  Query Match             16.7%;  Score 493;  DB 2;  Length 552;
  Best Local Similarity   28.6%;  Pred. No. 5.3e-26;
  Matches  159;  Conservative   76;  Mismatches  142;  Indels  178;  Gaps   26;

Qy      15 FKLWAPYVNSVKLKLS---KKLIPMEKNDEGFFE--VEIDDIEENL---TYSYIIEDK 64
           | |:|||     |  |    : |||:| |:|:|   ||::|         | |:  |:
Db       7 FTLFAPYNKGAALIASFSDWQEIPMKKGDDGYFRTTVELEDGTYQYKFRVQTRSWFFEED 66

Qy      65 R--EIPDPASR--YQPLGVHDKSQLIRTDYQILDLGKVKIED--------LIIYELHVGT 112
           :  :: || :    :  |  :    |:   :|:|    : :|        |:|||||||
Db      67 QWVDVTDPYATDIDESSGKDNSIARIKDGEKIVDTYVWQHDDKPLPADHELVIYELHVGD 126

Qy     113 FS-------QEGNFKGVIEKLDYLKDLGITGIELMPVAQFPGNRDWGYDGVFLYAVQNTY 165
           ||         | :| |||||||| :|||  |||:|| ::||:  |||:  : :| :::|
Db     127 FSGGEDDPYARGKYKHVIEKLDYLCELGINAIELLPVKEYPGDYSWGYNPRYFFATESSY 186

Qy     166 GGPWELAKLVNEAHKRGIAVILDVVYNHI--------------------GPEGNYLLGLG 205
           |   :| |||:| |:||| :|:| :|||                     |: |
Db     187 GSTADLKKLVDECHQRGIRIIMDGIYNHSEASSPLTQIDHDYWYHHEPRDPDNN------ 240

Qy     206 PYFSDRYKTPWGLTFNFD--DRGCD--QVRKFILENVEYWFKTFKIDGLRLDAVHAIFDN 261
                 ||  ||::  |   :    |||| : | ||   : :||:| ||   | :
Db     241 ----------WGPEFNYEHYDENLETYPARKFIGDTVRYWVGEYHLDGIRYDAARQIANY 290

Qy     262 SPKH-ILQEIAEKAHQLGKFVIAE--------SDLNDPKIVKDDCGYKIDAQWVDDFHHA 312
              | | ||  : |     : :||        ::|: |          :|  | | |:|
Db     291 DFMHWIAQEAKKTAGAKPFYNVAEHIPETTSITNLDGP----------MDGCWHDSFYHT 340

Qy     313 VHAFITKEKDYYYQDFGRIEDIEKTFKDVFVYDGKYSRYRGRTHGAPVGDLPPRKFVVFI 372
           : | |          |  |:|  |||  | |   : | |:          | ::
Db     341 IKAHIC----------GDTFDLE-NLKDVI--DPKRQGFLGATN----------VVNYL 376

Qy     373 QNHD------QVGNRGNGERLSILTDK--------TTYLMAATLYILSPYIPLIFMGEEY 418
            |||      ::|||       |  |:        |  || |        :|||:||||:
Db     377 TNHDHDHIMVELGNR------EIFHDEAFRRAKLGTAILMTAV------GVPLIWMGEEF 424

Qy     419 YETNPFFFFSDFSDPVLIKGVREGRLKENNQMIDPQSEEAFLKSKLSW-----KIDEEVL 473
           |  |`                    |: :|            ||: |      ::  :
Db     425 GEYKP--------------------KQQDQ-----------SKIDWTLLGNDLNRSLF 451

Qy     474 DYYKQLINIRKRYNN 488
           ||:| || :||  ||
Db     452 DYHKGLIGLRK--NN 464


Search completed: July 15, 2002, 12:18:44
Job time: 136 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: July 15, 2002, 12:24:32 ; Search time 19.24 Seconds
(without alignments)
1118.923 Million cell updates/sec

Title: US-09-298-924-8
Perfect score: 2954
Sequence: 1 MFSFGGNIEKNKGIFKLWAP.....KLKKDELIKVNRGVGVYQLE 556

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	969	32.8	596	1	TREZ_RHISP	Q53238 rhizobium s
2	959	32.5	598	1	TREZ_ARTSP	Q44316 arthrobacte
3	936	31.7	575	1	TREZ_ARTRM	Q9ajn6 arthrobacte
4	890.5	30.1	589	1	TREZ_BREHE	O52520 brevibacter
5	866.5	29.3	580	1	TREZ_MYCTU	Q10769 mycobacteri
6	353	11.9	666	1	GLGB_BACCL	P30537 bacillus ca
7	342.5	11.6	639	1	GLGB_BACST	P30538 bacillus st
8	320	10.8	843	1	PULA_THEMA	O33840 thermotoga
9	315	10.7	562	1	AMY2_DICTH	P14898 dictyoglomu
10	313.5	10.6	627	1	GLGB_BACSU	P39118 bacillus su
11	296.5	10.0	764	1	GLGB_STRAU	P52980 streptomyce
12	282	9.5	770	1	GLGB_SYNY3	P52981 synechocyst
13	276.5	9.4	659	1	GLGX_HAEIN	P45178 haemophilus
14	273	9.2	731	1	GLGB_MYCTU	Q10625 mycobacteri
15	270	9.1	702	1	GLGB_HUMAN	Q04446 homo sapien
16	269	9.1	734	1	GLGB_AGRTU	P52979 agrobacteri
17	267	9.0	639	1	GLGB_BUTFI	P30539 butyrivibri
18	266	9.0	730	1	GLGB_HAEIN	P45177 haemophilus
19	265	9.0	520	1	AMY_BACME	P20845 bacillus me
20	264.5	9.0	773	1	GLGB_SYNP7	P16954 synechococc
21	262.5	8.9	498	1	AMY3_DICTH	P14899 dictyoglomu
22	261.5	8.9	558	1	O16G_BACCE	P21332 bacillus ce
23	257	8.7	561	1	TREC_BACSU	P39795 bacillus su
24	254.5	8.6	562	1	O16G_BACTR	P29094 bacillus th
25	252.5	8.5	512	1	AMY1_DEBOC	P19269 debaryomyce
26	250.5	8.5	1196	1	AMYB_PAEPO	P21543 paenibacill
27	249	8.4	588	1	NEPU_BACST	P38940 bacillus st
28	243.5	8.2	478	1	YDD1_SCHPO	Q10427 schizosacch
29	243	8.2	561	1	O16G_BACSU	O06994 bacillus su
30	239.5	8.1	574	1	CDAS_THEET	P29964 thermoanaer
31	237	8.0	494	1	AMY1_SACFI	P21567 saccharomyc
32	234	7.9	555	1	O16G_BACCO	Q45101 bacillus co
33	234	7.9	1475	1	APU_THETY	P16950 t amylopull
34	231	7.8	1481	1	APU_THEET	P38939 t amylopull
35	229.5	7.8	508	1	O16G_BACSP	P29093 bacillus sp
36	228	7.7	799	1	GLGB_MAIZE	Q08047 zea mays (m
37	225.5	7.6	498	1	AMYA_ASPAW	Q02905 aspergillus
38	225.5	7.6	499	1	AMYA_ASPOR	P10529 aspergillus
39	225.5	7.6	499	1	AMYB_ASPAW	Q02906 aspergillus
40	225.5	7.6	569	1	MALT_CANAL	Q02751 candida alb
41	224.5	7.6	535	1	DEXB_STRPN	Q54796 streptococc
42	224	7.6	586	1	AMYM_BACAD	P32818 bacillus ac
43	224	7.6	728	1	GLGB_ECOLI	P07762 escherichia
44	223.5	7.6	499	1	AMY_ASPSH	P30292 aspergillus
45	223.5	7.6	777	1	ISOA_FLASP	O32611 flavobacter

ALIGNMENTS

RESULT 1
TREZ_RHISP
ID TREZ_RHISP STANDARD; PRT; 596 AA.
AC Q53238;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Malto-oligosyltrehalose trehalohydrolase (EC 3.2.1.141) (MTHase) (4-
DE alpha-D-{(1->4)-alpha-D-glucano}trehalose trehalohydrolase)
DE (Maltooligosyl trehalose trehalohydrolase).
GN TREZ.
OS Rhizobium sp. (strain M-11).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=391;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96219094; PubMed=8829547;
RA Maruta K., Hattori K., Nakada T., Kubota M., Sugimoto T., Kurimoto M.;
RT "Cloning and sequencing of trehalose biosynthesis genes from Rhizobium
RT sp. M-11.";
RL Biosci. Biotechnol. Biochem. 60:717-720(1996).
CC -!- CATALYTIC ACTIVITY: Hydrolysis of alpha-(1->4)-D-glucosidic
CC linkage in 4-alpha-D-{(1->4)-alpha-D-glucanosyl}(n) trehalose to
CC yield trehalose and alpha-(1->4)-D-glucan.
CC -!- PATHWAY: Trehalose biosynthesis.
CC -!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC --

CC --
DR EMBL; D78001; BAA11187.1; -.
DR InterPro; IPR000461; Alpha_amylase.
DR Pfam; PF00128; alpha-amylase; 1.
KW Hydrolase; Glycosidase.
FT ACT_SITE 265 265 BY SIMILARITY.
SQ SEQUENCE 596 AA; 65262 MW; 73EA80AE0534DDCD CRC64;

Query Match 32.8%; Score 969; DB 1; Length 596;
Best Local Similarity 39.5%; Pred. No. 1e-60;
Matches 215; Conservative 74; Mismatches 195; Indels 60; Gaps 12;

```
Qy      13 GIFKLWAPYVNSVKLKLSKKLIPMEKN------DEGFFEVEIDDIEENLTYSYIIE-DKR 65
           | | :|||   :| |    :   | :       |||::         :: | |::: |:
Db      15 GRFDIWAPEAGTVTLLAGGERYEMGRRPGNGPADEGWWTAADAPTGADVDYGYLLDGDEI 74

Qy      66 EIPDPASRYQPLGVHDKSQLIRTDYQILDLGKVKIEDL----------IIYELHVGTFSQ 115
              :||| :| || |||    | ||     | |  : :|          :|||||:|||:
Db      75 PLPDPRTRRQPEGVH---ALSRT----FDPGAHRWQDAGWQGRELQGSVIYELHIGTFTP 127
```

```
Qy    116 EGNFKGVIEKLDYLKDLGITGIELMPVAQFPGNRDWGYDGVFLYAVQNTYGGPWELAKLV 175
          ||       ||||| |||  |||:||  | |  :||||||  :||   ||||    : |
Db    128 EGTLDAAAGKLDYLAGLGIDFIELLPVNAFNGTHNWGYDGVQWFAVHEGYGGPAAYQRFV 187

Qy    176 NEAHKRGIAVILDVVYNHIGPEGNYLLGLGPYFSDRYKTPWGLTFNFDDRGCDQVRKFIL 235
          : ||  |: || ||||||:|| ||||   |||         || : | |  | | ||::||
Db    188 DAAHAAGLGVIQDVVYNHLGPSGNYLPRYGPYLKHGEGNTWGDSVNLDGPGSDHVRQYIL 247

Qy    236 ENVEYWFKTFKIDGLRLDAVHAIFDNSPKHILQEIAEKAHQLGK------FVIAESDLND 289
          :||  | : :::|||||||||||: |    |||:|   |  |         :|||||||:
Db    248 DNVAMWLRDYRVDGLRLDAVHALKDERAVHILEEFGALADALSSEGGRPLTLIAESDLNN 307

Qy    290 PKIV--KDDCGYKIDAQWVDDFHHAVHAFITKEKDYYYQDFGRIEDIEKTFKDVFVYDGK 347
          |:::  :|  || :  || |||||||  :: |   || ||  :  : |  :| | :||
Db    308 PRLLYPRDVNGYGLAGQWSDDFHHAVHVNVSGETTGYYSDFDSLGALAKVLRDGFFHDGS 367

Qy    348 YSRYRGRTHGAPV--GDLPPRKFVVFIQNHDQVGNRGNGERLSILTDKTTYLMAATLYIL 405
          || :||| || |:    : |   ||  |||||:|||  |:|||       :  :|| | :
Db    368 YSSFRGRCHGRPINFSAVHPAALVVCSQNHDQIGNRATGDRLSQSLPYGSLALAAVLTLT 427

Qy    406 SPYIRLIFMGEEYYETNPFFFFSDFSDPVLIKGVREGRLKENNQM-------IDPQSEEA 458
          |: |::|||||  | |: ||:   :| | |   |||::|  :|       |||  |
Db    428 GPFTPMLFMGEEYGATTPWQFFTSHPEPELGKATAEGRIREFERMGWDPAVVPDPQDPET 487

Qy    459 FLKSKLSWK-----IDEEVLDYYKQLINIRKRYNNCKRVKEVRREGNCITLIMEKIGIIA 513
          | :||| |          :|: |: || :|      :   |: | |   | :
Db    488 FTRSKLDWAEASAGDHARLLELYRSLITLR------RSTPELARLGFADTAV-------- 533

Qy    514 SFDD 517
            |||
Db    534 EFDD 537


RESULT   2
TREZ_ARTSP
ID   TREZ_ARTSP     STANDARD;      PRT;   598 AA.
AC   Q44316;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Malto-oligosyltrehalose trehalohydrolase (EC 3.2.1.141) (MTHase) (4-
DE   alpha-D-{(1->4)-alpha-D-glucano}trehalose trehalohydrolase)
DE   (Maltooligosyl trehalose trehalohydrolase).
GN   TREZ.
OS   Arthrobacter sp. (strain Q36).
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Micrococcineae; Micrococcaceae; Arthrobacter.
OX   NCBI_TaxID=1667;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96195835; PubMed=8605217;
RA   Maruta K., Hattori K., Nakada T., Kubota M., Sugimoto T., Kurimoto M.;
RT   "Cloning and sequencing of trehalose biosynthesis genes from
RT   Arthrobacter sp. Q36.";
RL   Biochim. Biophys. Acta 1289:10-13(1996).
CC   -!- CATALYTIC ACTIVITY: Hydrolysis of alpha-(1->4)-D-glucosidic
CC       linkage in 4-alpha-D-{(1->4)-alpha-D-glucanosyl}n trehalose to
CC       yield trehalose and alpha-(1->4)-D-glucan.
CC   -!- PATHWAY: Trehalose biosynthesis.
CC   -!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC       KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D63343; BAA09668.1; -.
DR   InterPro; IPR000461; Alpha_amylase.
DR   InterPro; IPR004193; isoamylase_N.
DR   Pfam; PF00128; alpha-amylase; 1.
DR   Pfam; PF02922; isoamylase_N; 1.
KW   Hydrolase; Glycosidase.
FT   ACT_SITE    267    267       BY SIMILARITY.
SQ   SEQUENCE   598 AA;  65831 MW;  8B5C610AD3766947 CRC64;


  Query Match             32.5%;  Score 959;  DB 1;  Length 598;
  Best Local Similarity   40.2%;  Pred. No. 5.1e-60;
  Matches  202;  Conservative   81;  Mismatches  186;  Indels   34;  Gaps   10;

Qy     15 FKLWAPYVNSVKLKLSKKLIPMEK------NDEGFFEVEIDDIEENLTYSYIIE-DKREI 67
          : :|||   || |    :   |::       | |::       : |: | |::: |:  :
Db     19 YDVWAPNAESVTLLAGGERYAMQRRAETGPEDAGWWTAAGAPTDGNVDYGYLLDGDETPL 78

Qy     68 PDPASRYQPLGVHDKSQLI-RTDYQILD---LGKVKIEDLIIYELHVGTFSQEGNFKGVI 123
          ||| :| || ||| |:    : |   |    |: :::  :|||||:|||: ||  :
Db     79 PDPRTRRQPDGVHALSRTFDPSAYSWQDDAWQGR-ELQGAVIYELHLGTFTPEGTLEAAA 137

Qy    124 EKLDYLKDLGITGIELMPVAQFPGNRDWGYDGVFLYAVQNTYGGPWELAKLVNEAHKRGI 183
          |||||  ||:  |||:||  | |  :||||||  :||   ||||    : |: ||  |:
Db    138 GKLDYLAGLGVDFIELLPVNAFNGTHNWGYDGVQWFAVHEAYGGPEAYQRFVDAAHAAGL 197

Qy    184 AVILDVVYNHIGPEGNYLLGLGPYFSDRYKTPWGLTFNFDDRGCDQVRKFILENVEYWFK 243
          || ||||||:|| ||||   |||         || : | |  | | ||::||:|:  | :
Db    198 GVIQDVVYNHLGPSGNYLPRFGPYLKQGEGNTWGDSVNLDGPGSDHVRRYILDNLAMWLR 257

Qy    244 TFKIDGLRLDAVHAIFDNSPKHILQEIAEKAHQLGKFV------IAESDLNDPKIV--KD 295
          :::|||||||||||: |    |||::    | |:   |      |||||||:|:::  :|
Db    258 DYRVDGLRLDAVHALKDERAVHILEDFGALADQISAEVGRPLTLIAESDLNNPRLLYPRD 317

Qy    296 DCGYKIDAQWVDDFHHAVHAFITKEKDYYYQDFGRIEDIEKTFKDVFVYDGKYSRYRGRT 355
            || :: || ||||||||  :| |   || ||  :  : |  :| | :|| || :| |
Db    318 VNGYGLEGQWSDDFHHAVHVNVTGETTGYYSDFDSLAALAKVLRDGFFHDGSYSSFRERH 377

Qy    356 HGAPV--GDLPPRKFVVFIQNHDQVGNRGNGERLSILTDKTTYLMAATLYILSPYIPLIF 413
          || |:    : |   ||  |||||:|||  |:|||       :  :|| | :  |: |::
Db    378 HGRPINFSAVHPAALVVCSQNHDQIGNRATGDRLSQTLPYGSLALAAVLTLTGPFTPMLL 437

Qy    414 MGEEYYETNPFFFFSDFSDPVLIKGVREGRLKENNQM-------IDPQSEEAFLKSKLSW 466
          |||||  : |: ||:   :| | |   |||:||  :|        |||  | | :||| |
Db    438 MGEEYGASTPWQFFTSHPEPELGKATAEGRIKEFERMGWDPAVVPDPQDPETFRRSKLDW 497

Qy    467 KIDEE-----VLDYYKQLINIRK 484
            |      :|: |: |  :|:
Db    498 AEAAEGDHARLLELYRSLTALRR 520


RESULT   3
TREZ_ARTRM
ID   TREZ_ARTRM     STANDARD;      PRT;   575 AA.
AC   Q9AJN6;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Malto-oligosyltrehalose trehalohydrolase (EC 3.2.1.141) (MTHase) (4-
DE   alpha-D-[(1->4)-alpha-D-glucano]trehalose trehalohydrolase)
DE   (Maltooligosyl trehalose trehalohydrolase).
GN   TREZ.
OS   Arthrobacter ramosus.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Micrococcineae; Micrococcaceae; Arthrobacter.
OX   NCBI_TaxID=1672;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S34;
RA   Yamamoto T., Maruta K., Watanabe H., Yamashita H., Kubota M.,
RA   Fukuda S., Kurimoto M.;
RT   "Trehalose producing operon treYZ from Arthrobacter ramosus S34.";
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- CATALYTIC ACTIVITY: Hydrolysis of alpha-(1->4)-D-glucosidic
CC       linkage in 4-alpha-D-{(1->4)-alpha-D-glucanosyl}(n) trehalose to
```

```
CC       yield trehalose and alpha-(1->4)-D-glucan.
CC   -!- PATHWAY: Trehalose biosynthesis.
CC   -!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC       KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AB045141; BAB40766.1; -.
DR   InterPro; IPR000461; Alpha_amylase.
DR   Pfam; PF00128; alpha-amylase; 1.
KW   Hydrolase; Glycosidase.
FT   ACT_SITE    250    250       BY SIMILARITY.
SQ   SEQUENCE   575 AA;  63079 MW;  284A3F20207E228B CRC64;


  Query Match             31.7%;  Score 936;  DB 1;  Length 575;
  Best Local Similarity   38.7%;  Pred. No. 2e-58;
  Matches  198;  Conservative   82;  Mismatches  177;  Indels   54;  Gaps   10;

Qy      15 FKLWAPYVNSVKLKL--SKKLIPMEKNDEGFFEVE--IDDIEENLTYSYIIEDKREIPDP 70
           | :|||    | | :   :  :|: ::: |:: ::   |   : : | |::: |    ||
Db       5 FPVWAPQAAQVTLVVGQGRAELPLTRDENGWWALQQPWDGGPDLVDYGYLVDGKGPFADP 64

Qy      71 ASRYQPLGVHDKSQLIRTDYQILDLGKV----------------KIEDLIIYELHVGTFS 114
             |  || |||              :||:              :   :||||||||||:
Db      65 RSLRQPRGVH-------------ELGREFDPARYAWGDDGWRGRDLTGAVIYELHVGTFT 111

Qy     115 QEGNFKGVIEKLDYLKDLGITGIELMPVAQFPGNRDWGYDGVFLYAVQNTYGGPWELAKL 174
           ||     | :||:|  ||:  :||:||  | |   ||||||  |||   ||||    :
Db     112 PEGTLDSAIRRLDHLVRLGVDAVELLPVNAFNGTHGWGYDGVLWYAVHEPYGGPEAYQRF 171

Qy     175 VNEAHKRGIAVILDVVYNHIGPEGNYLLGLGPYFSDRYKTPWGLTFNFDDRGCDQVRKFI 234
           |:  | ||:||: ||||||:|| ||:|   |||         ||   | |    |:||::|
Db     172 VDACHARGLAVVQDVVYNHLGPSGNHLPDFGPYLGSGAANTWGDALNLDGPLSDEVRRYI 231

Qy     235 LENVEYWFKTFKIDGLRLDAVHAIFDNSPKHILQEIA----EKAHQLGK--FVIAESDLN 288
           ::|  || :    |||||||||: |    |:|:|:|    | | :||:   :|||||||
Db     232 IDNAVYWLRDMHADGLRLDAVHALRDARALHLLEELAARVDELAGELGRPLTLIAESDLN 291

Qy     289 DPKIVKDDC--GYKIDAQWVDDFHHAVHAFITKEKDYYYQDFGRIEDIEKTFKDVFVYDG 346
           |||:::     || :|||| || |||||| :| |   || ||| :  : | |:  : :||
Db     292 DPKLIRSRAAHGYGLDAQWDDDVHHAVHANVTGETVGYYADFGGLGALVKVFQRGWFHDG 351

Qy     347 KYSRYRGRTHGAPVG-DLPPRKFVVFIQNHDQVGNRGNGERLSILTDKTTYLMAATLYIL 405
           :| :| | || |:  |:| |: | | |:|||||||  |:|:|    : :   || | :|
Db     352 TWSSFRERHHGRPLDPDIPFRRLVAFAQDHDQVGNRAVGDRMSAQVGEGSLAAAAALVLL 411

Qy     406 SPYIPLIFMGEEYYETNPFFFFSDFSDPVLIKGVREGRLKENNQM-------IDPQSEEA 458
             |: |::|||||:    |: ||:   :| | :    ||: |  :|       |||
Db     412 GPFTPMLFMGEEWGARTPWQFFTSHPEPELGEATARGRIAEFARMGWDPAVVPDPQDPAT 471

Qy     459 FLKSKLSWKIDEE-----VLDYYKQLINIRK 484
           | :| | |    |      :| :|  || :|:
Db     472 FARSHLDWSEPEREPHAGLLAFYTDLIALRR 502


RESULT   4
TREZ_BREHE
ID   TREZ_BREHE     STANDARD;      PRT;   589 AA.
AC   O52520;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Malto-oligosyltrehalose trehalohydrolase (EC 3.2.1.141) (MTHase) (4-
DE   alpha-D-{(1->4)-alpha-D-glucano}trehalose trehalohydrolase)
DE   (Maltooligosyl trehalose trehalohydrolase).
GN   TREZ.
OS   Brevibacterium helvolum.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Micrococcineae; Brevibacteriaceae; Brevibacterium.
OX   NCBI_TaxID=1704;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 11822;
RA   Kwon T., Kim C.H., Choi Y.D.;
RT   "Maltooligosyl trehalose synthase/trehalohydrolase encode proteins for
RT   trehalose production in Brevibacterium helvolum.";
RL   Submitted (DEC-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- CATALYTIC ACTIVITY: Hydrolysis of alpha-(1->4)-D-glucosidic
CC       linkage in 4-alpha-D-{(1->4)-alpha-D-glucanosyl}(n) trehalose to
CC       yield trehalose and alpha-(1->4)-D-glucan.
CC   -!- PATHWAY: Trehalose biosynthesis.
CC   -!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC       KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF039919; AAB95369.1; -.
DR   InterPro; IPR000461; Alpha_amylase.
DR   Pfam; PF00128; alpha-amylase; 1.
KW   Hydrolase; Glycosidase.
FT   ACT_SITE    258    258       BY SIMILARITY.
SQ   SEQUENCE   589 AA;  64217 MW;  C282314B6E9BA029 CRC64;


  Query Match             30.1%;  Score 890.5;  DB 1;  Length 589;
  Best Local Similarity   39.1%;  Pred. No. 3.2e-55;
  Matches  197;  Conservative   76;  Mismatches  198;  Indels   33;  Gaps   10;

Qy      15 FKLWAPYVNSVKLKLSKKLIPMEKND-----EGFFEVEIDDIEENLTYSYIIE-DKREIP 68
           | :||| |:|| |    :  ||:| :     ||::         :: | |::: :   :|
Db      11 FDVWAPDVSSVVLVADGRQYPMQKKETAPGSEGWWTASDAPPNGDVDYGYLLDGNTTPVP 70

Qy      69 DPASRYQPLGVHDKSQLIR-TDYQILD---LGKVKIEDLIIYELHVGTFSQEGNFKGVIE 124
           :| ||  | |||: |:      |:  |    || :::  :||:|||| : :|      |
Db      71 EPRSRRLPAGVHNHSRTYNPPPYRWQDSRWRGK-ELQGTLIYQLHVGTSTPDGTLDAAGE 129

Qy     125 KLDYLKDLGITGIELMPVAQFPGNRDWGYDGVFLYAVQNTYGGPWELAKLVNEAHKRGIA 184
           || || ||||  |||:||  | |  :||||||  | |   ||||    : |: ||  |:
Db     130 KLSYLVDLGIDFIELLPVNGFNGTHNWGYDGVQWYTVHEGYGGPAAYQRFVDAAHAAGLG 189

Qy     185 VILDVVYNHIGPEGNYLLGLGPYFSDRYKTPWGLTFNFDDRGCDQVRKFILENVEYWFKT 244
           || ||||||:|  |||   |||          | : | |  | |  |::||:|   |
Db     190 VIQDVVYNHLGLRGNYFPKLGPNLKQGDANTLGDSVNLDGAGSDVFREYILDNAALWVGD 249

Qy     245 FKIDGLRLDAVHAIFDNSPKHILQEIAEKAHQ------LGKFVIAESDLNDPKIV--KDD 296
           : :||:  |||||: |    |||:::             | | :|||||  |:|:::  :|
Db     250 YHVDGVGFDAVHAVRDERAVHILEDLGALGDAISGETGLPKTLIAESDFNNPRLIYPRDV 309

Qy     297 CGYKIDAQWVDDFHHAVHAFITKEKDYYYQDFGRIEDIEKTFKDVFVYDGKYSRYRGRTH 356
            || :  || |||| |||  :: |   || ||  :  : |  || |::|| || :||| |
Db     310 NGYGLAGQWSDDFHTAVHVSVSGETTGYYSDFESLAVLAKVLKDGFLHDGSYSSFRGRHH 369

Qy     357 GAPVGD--LPPRKFVVFIQNHDQVGNRGNGERLSILTDKTTYLMAATLYILSPYIPLIFM 414
           | |:      |   ||  |||||:|||  |:|||       :|| | : ||: |::||
Db     370 GRPINPSLANPAALVVCNQNHDQIGNRATGDRLSQSLSYGQLAVAAVLTLTSPFTPMLFM 429

Qy     415 GEEYYETNPFFFFSDFSDPVLIKGVREGRLKENNQM-------IDPQSEEAFLKSKLSWK 467
           ||||  : |: ||:   :| | |   |||:||  :|        |||  | | :||| |
Db     430 GEEYGASTPWQFFTSHPEPELGKATAEGRIKEFERMGWDPAVVPDPQDPETFNRSKLDWS 489

Qy     468 -----IDEEVLDYYKQLINIRKRY 486
                :|: || |  :|: :
Db     490 EASTGDHARLLELYKSLTALRREH 513
```

```
RESULT   5
TREZ_MYCTU
ID   TREZ_MYCTU     STANDARD;      PRT;   580 AA.
AC   Q10769;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Malto-oligosyltrehalose trehalohydrolase (EC 3.2.1.141) (MTHase) (4-
DE   alpha-D-[(1->4)-alpha-D-glucano]trehalose trehalohydrolase)
DE   (Maltooligosyl trehalose trehalohydrolase).
GN   TREZ OR RV1562C OR MT1613 OR MTCY48.03.
OS   Mycobacterium tuberculosis.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX   NCBI_TaxID=1773;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=H37RV;
RX   MEDLINE=98295987; PubMed=9634230;
RA   Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA   Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekaia F.,
RA   Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA   Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA   Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA   Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA   Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA   Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT   "Deciphering the biology of Mycobacterium tuberculosis from the
RT   complete genome sequence.";
RL   Nature 393:537-544(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CDC 1551 / Oshkosh;
RA   Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA   Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA   Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA   Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA   Bishai W.;
RT   "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT   laboratory strains.";
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- CATALYTIC ACTIVITY: Hydrolysis of alpha-(1->4)-D-glucosidic
CC       linkage in 4-alpha-D-[(1->4)-alpha-D-glucanosyl](n) trehalose to
CC       yield trehalose and alpha-(1->4)-D-glucan.
CC   -!- PATHWAY: Trehalose biosynthesis.
CC   -!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC       KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z74020; CAA98329.1; -.
DR   EMBL; AE007027; AAK45880.1; ALT_INIT.
DR   TIGR; MT1613; -.
DR   TubercuList; Rv1562c; -.
DR   InterPro; IPR000461; Alpha_amylase.
DR   InterPro; IPR004193; isoamylase_N.
DR   Pfam; PF00128; alpha-amylase; 1.
DR   Pfam; PF02922; isoamylase_N; 1.
KW   Hydrolase; Glycosidase; Complete proteome.
FT   ACT_SITE    247    247       BY SIMILARITY.
SQ   SEQUENCE   580 AA;  64076 MW;  506190468F44B862 CRC64;


  Query Match             29.3%;  Score 866.5;  DB 1;  Length 580;
  Best Local Similarity   37.3%;  Pred. No. 1.5e-53;

  Matches  187;  Conservative   83;  Mismatches  194;  Indels   37;  Gaps   11;

Qy      15 FKLWAPYVNSVKLKLSKKLIPMEKNDEGFFEVEIDDIEENLTYSYIIEDKREI-PDPASR 73
           |::|||    |:| ::  :  | :: :|::   :      :  | |:::|   : ||| |
Db       4 FRVWAPKPALVRLDVNGAVHAMTRSADGWWHTTV-AAPADARYGYLLDDDPTVLPDPRSA 62

Qy      74 YQPLGVHDKSQ---------LIRTDYQILDLGKVKIEDLIIYELHVGTFSQEGNFKGVIE 124
             || ||| :||          |||          :|  :|||||:|||:  | |   ||
Db      63 RQPDGVHARSQRWEPPGQFGAARTD---TGWPGRSVEGAVIYELHIGTFTTAGTFDAAIE 119

Qy     125 KLDYLKDLGITGIELMPVAQFPGNRDWGYDGVFLYAVQNTYGGPWELAKLVNEAHKRGIA 184
           ||||| |||| :|||||  | | | ||||||  |:|   ||||  | : ::  | | :
Db     120 KLDYLVDLGIDFVELMPVNSFAGTRGWGYDGVLWYSVHEPYGGPDGLVRFIDACHARRLG 179

Qy     185 VILDVVYNHIGPEGNYLLGLGPYFSDRYKTPWGLTFNFDDRGCDQVRKFILENVEYWFKT 244
           |::| |:||:|| ||||   ||| |     |||   |     |:|| :|::   | :
Db     180 VLIDAVFNHLGPSGNYLPRFGPYLSSA-SNPWGDGINIAGADSDEVRHYIIDCALRWMRD 238

Qy     245 FKIDGLRLDAVHAIFDNSPKHILQEIAEKAH----QLGK--FVIAESDLNDPKIV--KDD 296
           |  ||||||||||: | :  |:|:|:|        |||:   :|||:| |||:::
Db     239 FHADGLRLDAVHALVDTTAVHVLEELANATRWLSGQLGRPLSLIAETDRNDPRLITRPSH 298

Qy     297 CGYKIDAQWVDDFHHAVHAFITKEKDYYYQDFGRIEDIEKTFKDVFVYDGKYSRYRGRTH 356
           || | ||| || |||:|  :: |:  || ||| :  :  | :: : : | ||':| | |
Db     299 GGYGITAQWNDDIHHAIHTAVSGERQGYYADFGSLATLAYTLRNGYFHAGTYSSFRRRRH 358

Qy     357 GAPV--GDLPPRKFVVFIQNHDQVGNRGNGERLSILTDKTTYLMAATLYILSPYIPLIFM 414
           |  :    :|  : : :   |||||||  |:| |          : | | : |||  ::||
Db     359 GRALDTSAIPATRLLAYTCTHDQVGNRALGDRPSQYLTGGQLAIKAALTLGSPYTAMLFM 418

Qy     415 GEEYYETNPFFFFSDFSDPVLIKGVREGRLKE-------NNQMIDPQSEEAFLKSKLSWK 467
           |||:  ::|| ||    :| |      || :|        : : |||  : | : ||:|
Db     419 GEEWGASSPFQFFCSHPEPELAHSTVAGRKEEFAEHGWAADDIPDPQDPQTFQRCKLNWA 478

Qy     468 -----IDEEVLDYYKQLINIR 483
                 :  :|: || :|
Db     479 EAGSGEHARLHRFYRDLIALR 499


RESULT   6
GLGB_BACCL
ID   GLGB_BACCL     STANDARD;      PRT;   666 AA.
AC   P30537;
DT   01-APR-1993 (Rel. 25, Created)
DT   01-APR-1993 (Rel. 25, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   1,4-alpha-glucan branching enzyme (EC 2.4.1.18) (Glycogen branching
DE   enzyme).
GN   GLGB.
OS   Bacillus caldolyticus.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Bacillus/Staphylococcus group; Bacillus.
OX   NCBI_TaxID=1394;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93208370; PubMed=1296817;
RA   Kiel J.A.K.W., Boels J.M., Beldman G., Venema G.;
RT   "The glgB gene from the thermophile Bacillus caldolyticus encodes a
RT   thermolabile branching enzyme.";
RL   DNA Seq. 3:221-232(1992).
CC   -!- CATALYTIC ACTIVITY: Formation of 1,6-glucosidic linkages of
CC       glycogen.
CC   -!- PATHWAY: THIRD STEP IN GLYCOGEN BIOSYNTHESIS.
CC   -!- MISCELLANEOUS: OPTIMAL ACTIVITY AT APPROXIMATELY 39 DEGREES
CC       CELSIUS.
CC   -!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC       KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
```

```
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; Z14057; CAA78440.1; -.
DR   InterPro; IPR000461; Alpha_amylase.
DR   InterPro; IPR004193; isoamylase_N.
DR   Pfam; PF00128; alpha-amylase; 1.
DR   Pfam; PF02922; isoamylase_N; 1.
KW   Glycogen biosynthesis; Transferase; Glycosyltransferase.
FT   ACT_SITE    309    309       BY SIMILARITY.
FT   ACT_SITE    352    352       BY SIMILARITY.
FT   ACT_SITE    420    420       BY SIMILARITY.
SQ   SEQUENCE   666 AA;  78096 MW;  1F64C31495BB1B6E CRC64;


  Query Match             11.9%;  Score 353;  DB 1;  Length 666;
  Best Local Similarity   23.9%;  Pred. No. 2.2e-17;
  Matches  142;  Conservative   95;  Mismatches  192;  Indels  166;  Gaps   29;

Qy       4 FGGNIEKNKGI----FKLWAPYVNSVKL-----KLSKKLIPMEK-NDEGFFEVEIDDIEE 53
           || ::  :  |     | :|||:   |:|       :    |: | |||| : : : :  |
Db      26 FGAHVIRGGGAVGTRFCVWAPHAREVRLVGSFNDWNGTNSPLTKVNDEGVWTIVVPENLE 85

Qy      54 NLTYSY--IIEDKREI--PDPASRYQPLGVHDKS---------------QLIRTDYQILD 94
             | |  |  | |  :   || : |  |  |  |                |  :   :| |
Db      86 GHLYKYEIITPDGRVLLKADPYAFYSELRPHTASIVYDLKGYEWNDSPWQRKKRRKRIYD 145

Qy      95 LGKVKIEDLIIYELHVGTFSQ--EGNFKGVIEKLD----YLKDLGITGIELMPVAQFPGN 148
             : ::|||||  |:: :  :| |    |  |    |: : | | |||:|: : | :
Db     146 ------QPMVIYELHFGSWKKKPDGRFYTYREMADELIPYVLERGFTHIELLPLVEHPLD 199

Qy     149 RDWGYDGVFLYAVQNTYGGPWELAKLVNEAHKRGIAVILDVVYNHI--GPEGNYLLGLGP 206
           | ||| |   |:| : || | :    |:  |: |: ||:| |  |      | |:    |
Db     200 RSWGYQGTGYYSVTSRYGTPHDFMYFVDRCHQAGLGVIIDWVPGHFCKDAHGLYMFDGAP 259

Qy     207 YF-----SDRYKTPWGLTFNFDDRGCDQVRKFILENVEYWFKTFKIDGLRLDAV------ 255
            :       ||    || | || | |  :|| |:: |  :| : : :|| |:|||
Db     260 TYEYANEKDRENYVWG-TANF-DLGKPEVRSFLISNALFWLEYYHVDGFRVDAVANMLYW 317

Qy     256 ---HAIFDNS-PKHILQEIAEK--AHQLGKFVIAESDLNDPKIVKD--DCG----YKIDA 303
               `:::|      |::: |   |:    ::|||   : |::     | |    || :
Db     318 PNNDRLYENPYAVEFLRQLNEAVFAYDPNVWMIAEDSTDWPRVTAPTYDGGLGFNYKWNM 377

Qy     304 QWVDDFHHAVHAFITKEKDYYYQDFGRIEDIEKTFKDVFVYDGKYSRYRGRTHGAPVGDL 363
           |::|    :      |: | :         : :|  :: |   :
Db     378 GWMNDMLKYMET-PPHERKYAHN--------QVSFSLLYAYSENF--------------- 413

Qy     364 PPRKFVVFIQNHDQVG-------NRGNG---ERLSILTDKTTYLMAATLYILSPYIPLIF 413
           :   :||:|      |:  |   |: : |     |:||       |   |:|
Db     414 ------ILPFSHDEVVHGKKSLLNKMPGSYEEKFAQLRLLYGYMMA------HPGKKLLF 461

Qy     414 MGEEYYETNPFFFFSDFSDPVLIKGVREGRLKENNQMIDPQSEEAFLKSKLSWKIDEEVL 473
           || |: : : : |  :  | ||                          :|  |:||
Db     462 MGSEFAQFDEWKFAEEL-DWVL-------------------------FDFELHRKMDE--- 493

Qy     474 DYYKQLINIRKRY-----------------NNCKR-----VKEVRREGNCITLI 505
           | ||||   |||                  :| ::     ::  ::||: : ::
Db     494 -YVKQLIACYKRYKPFYELDHDPRGFEWIDVHNAEQSIFSFIRRGKKEGDVLVIV 547


RESULT   7
GLGB_BACST
ID   GLGB_BACST     STANDARD;      PRT;   639 AA.
AC   P30538;
DT   01-APR-1993 (Rel. 25, Created)
DT   01-APR-1993 (Rel. 25, Last sequence update)
DT   01-FEB-1995 (Rel. 31, Last annotation update)
DE   1,4-alpha-glucan branching enzyme (EC 2.4.1.18) (Glycogen branching
DE   enzyme).
GN   GLGB.
OS   Bacillus stearothermophilus.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Bacillus/Staphylococcus group; Geobacillus.
OX   NCBI_TaxID=1422;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=1503-4R;
RX   MEDLINE=92079888; PubMed=1745226;
RA   Kiel J.A.K.W., Boels J.M., Beldman G., Venema G.;
RT   "Molecular cloning and nucleotide sequence of the glycogen branching
RT   enzyme gene (glgB) from Bacillus stearothermophilus and expression in
RT   Escherichia coli and Bacillus subtilis.";
RL   Mol. Gen. Genet. 230:136-144(1991).
CC   -!- CATALYTIC ACTIVITY: Formation of 1,6-glucosidic linkages of
CC       glycogen.
CC   -!- PATHWAY: THIRD STEP IN GLYCOGEN BIOSYNTHESIS.
CC   -!- MISCELLANEOUS: THE TEMPERATURE FOR OPTIMAL ACTIVITY IS
CC       APPROXIMATELY 55 DEGREES CELSIUS.
CC   -!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC       KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; M35089; AAA22482.1; -.
DR   PIR; S18599; S18599.
DR   InterPro; IPR000461; Alpha_amylase.
DR   InterPro; IPR004193; isoamylase_N.
DR   Pfam; PF00128; alpha-amylase; 1.
DR   Pfam; PF02922; isoamylase_N; 1.
KW   Glycogen biosynthesis; Transferase; Glycosyltransferase.
FT   ACT_SITE    309    309       BY SIMILARITY.
FT   ACT_SITE    352    352       BY SIMILARITY.
FT   ACT_SITE    420    420       BY SIMILARITY.
SQ   SEQUENCE   639 AA;  74795 MW;  F5BD4446B371E03A CRC64;


  Query Match             11.6%;  Score 342.5;  DB 1;  Length 639;
  Best Local Similarity   24.9%;  Pred. No. 1.1e-16;
  Matches  135;  Conservative   89;  Mismatches  197;  Indels  121;  Gaps   26;

Qy       4 FGGNIEKNKGI----FKLWAPYVNSVKLKLS------KKLIPMEKNDEGFFEVEIDDIEE 53
           || ::  |  |:    | :|||:   |:|  |           |: :::| : : | :  |
Db      26 FGAHVIKKNGMVGTRFCVWAPHAREVRLVGSFNEWNGTNFNLMKVSNQGVWMIFIPENLE 85

Qy      54 NLTYSYII--EDKREI--PDPASRYQPLGVHDKSQLIR-TDYQILD--------LGKVKI 100
             | |  |   |   :   || : |  |  |  | :     ||  |            ::
Db      86 GHLYKYEITTNDGNVLLKSDPYAFYSELRPHTASIVYNIKGYQWNDQTWRRKKQRKRIYD 145

Qy     101 EDLIIYELHVGTF--SQEGNF----KGVIEKLDYLKDLGITGIELMPVAQFPGNRDWGYD 154
           : | |||||  |::   ::|:|    :   | : |: : | | |||:|: : | :| |||
Db     146 QPLFIYELHFGSWKKKEDGSFYTYQEMAEELIPYVLEHGFTHIELLPLVEHPFDRSWGYQ 205

Qy     155 GVFLYAVQNTYGGPWELAKLVNEAHKRGIAVILDVVYNHI--GPEGNYLLGLGPYF---- 208
           |:  |:  : || | :|   ::  |: || |||| |  |      | |:    | :
Db     206 GIGYYSATSRYGTPHDLMYFIDRCHQAGIGVILDWVPGHFCKDSHGLYMFDGAPAYEYAN 265

Qy     209 -SDRYKTPWGLTFNFDDRGCDQVRKFILENVEYWFKTFKIDGLRLDAV---------HAI 258
              ||    || | || | |  :|| |:: |  :| : | :|| |:|||         :
Db     266 MQDRENYVWG-TANF-DLGKPEVRSFLISNALFWMEYFHVDGFRVDAVANMLYWPNSDVL 323

Qy     259 FDNS-PKHILQEIAEK--AHQLGKFVIAESDLNDPKIVKD--DCG----YKIDAQWVDDF 309
           : |:     ||:: |   |:     :|||   : |::     | |    || :  |::|
Db     324 YKNTYAVEFLQKLNETVFAYDPNILMIAEDSTDWPRVTAPTYDGGLGFNYKWNMGWMNDI 383

Qy     310 HHAVHAFITKEKDYYYQDFGRIEDIEKTFKDVFVYDGKYSRYRGRTHGAPVGDLPPRKFV 369
             :     |  | :       : ||  :: |   :
Db     384 --LTYMETPPEHRKYVHN-------KVTFSLLYAYSENF---------------------- 413

Qy     370 VFIQNHDQVGNRGNGERLSIL----TDKTTYLMAATLYILS-PYIPLIFMGEEYYETNPF 424
           :   :||:| : |    || :     :|   |     |:|: |   |:||| |: :
```

```
Db    414 ILPFSHDEVVH-GKKSLLSKMPGTYEEKFAQLRLLYGYLLTHPGKKLLFMGGEFGQ---- 468

Qy    425 FFFSDFSDPVLIKGVREGRLKENNQMIDPQSEEAFLKSKLSWKIDEEVLDYYKQLINIRK 484
          | :: |           |:: : |:             : :   :  | |:|:   |
Db    469 --FDEWKD----------LEQLDWML-----------FDFDMHRNMNMYVKELLKCYK 503

Qy    485 RY 486
          ||
Db    504 RY 505


RESULT   8
PULA_THEMA
ID   PULA_THEMA     STANDARD;      PRT;   843 AA.
AC   O33840;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Pullulanase precursor (EC 3.2.1.41) (Alpha-dextrin endo-1,6-alpha-
DE   glucosidase) (Pullulan 6-glucanohydrolase).
GN   PULA OR TM1845.
OS   Thermotoga maritima.
OC   Bacteria; Thermotogales; Thermotoga.
OX   NCBI_TaxID=2336;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MSB8 / DSM 3109;
RX   MEDLINE=98115241; PubMed=9453151;
RA   Bibel M., Brettl C., Gosslar U., Kriegshaeuser G., Liebl W.;
RT   "Isolation and analysis of genes for amylolytic enzymes of the
RT   hyperthermophilic bacterium Thermotoga maritima.";
RL   FEMS Microbiol. Lett. 158:9-15(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MSB8 / DSM 3109;
RX   MEDLINE=99287316; PubMed=10360571;
RA   Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
RA   Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA   McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
RA   Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
RA   Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA   Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT   "Evidence for lateral gene transfer between Archaea and Bacteria from
RT   genome sequence of Thermotoga maritima.";
RL   Nature 399:323-329(1999).
CC   -!- CATALYTIC ACTIVITY: Hydrolysis of (1->6)-alpha-D-glucosidic
CC       linkages in pullulan and in amylopectin and glycogen, and the
CC       alpha-and beta-limit dextrins of amylopectin and glycogen.
CC   -!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC       KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AJ001087; CAA04522.1; -.
DR   EMBL; AE001821; AAD36907.1; -.
DR   TIGR; TM1845; -.
DR   InterPro; IPR000461; Alpha_amylase.
DR   InterPro; IPR004193; isoamylase_N.
DR   Pfam; PF00128; alpha-amylase; 1.
DR   Pfam; PF02922; isoamylase_N; 1.
KW   Hydrolase; Glycosidase; Signal; Complete proteome.
FT   SIGNAL       1     19       POTENTIAL.
FT   CHAIN       20    843       PULLULANASE.
FT   ACT_SITE   535    535       BY SIMILARITY.
FT   ACT_SITE   564    564       BY SIMILARITY.
FT   ACT_SITE   652    652       BY SIMILARITY.
SQ   SEQUENCE   843 AA;  96261 MW;  C42DDE233D54FE77 CRC64;
```

```
  Query Match           10.8%;  Score 320;  DB 1;  Length 843;
  Best Local Similarity   23.4%;  Pred. No. 6.3e-15;
  Matches  143;  Conservative   88;  Mismatches  185;  Indels  194;  Gaps   31;

Qy      5 GGNIEKNKGIFKLWAPYVNSVKLKLSK--------KLIPMEKNDEGFFEVEIDDIEENLT 56
          |      | ||::|:|    ||: | |        ::: ||    | :|  ::   : :
Db    226 GAVYSPEKTIFRVWSPVSKWVKVLLFKNGEDTEPYQVVNMEYKGNGVWEAVVEGDLDGVF 285

Qy     57 YSYIIEDK---REIPDPASRYQPLGVHDKSQ------LIRTDYQ--ILDLG-KVK-IEDL 103
          | | :|:    |   || |:     |:  |:       | ||: :    | | |::  ||
Db    286 YLYQLENYGKIRTTVDPYSK----AVYANSKKSAVVNLARTNPEGWENDRGPKIEGYEDA 341

Qy    104 IIYELHVGTFS----------------QEGNFK---GVIEKLDYLKDLGITGIELMPVAQ 144
          ||||:|:   :                | | |   ||   | :| :||:| : ::|
Db    342 IIYEIHIADITGLENSGVKNKGLYLGLTEENTKGPGGVTTGLSHLVELGVTHVHILPFFD 401

Qy    145 F-PGNR---------DWGYDGVFLYAV---------QNTYGGPWELAKLVNEAHKRGIAV 185
          |  |:          :|||| :|: |          :| :    |: ::|   || || |
Db    402 FYTGDELDKDFEKYYNWGYD-PYLFMVPEGRYSTDPKNPHTRIREVKEMVKALHKHGIGV 460

Qy    186 ILDVVYNHIGPEGNYLLGLG---------PYFSDRY-KTPWGLTFNFDDRGCDQV----- 230
          |:|:|: |          |:|          ||:  |  ||  |   |  : ||  |
Db    461 IMDMVFPH-------TYGIGELSAFDQTVPYYFYRIDKT--GAYLN--ESGCGNVIASER 509

Qy    231 ---RKFILENVEYWFKTFKIDGLRLDAVHAIFDNSPKHILQEIAEKAHQLGKFVIAESDL 287
             ||||:: | || | : ||| | | :  |      |  : |:    |::
Db    510 PMMRKFIVDTVTYWVKEYHIDGFRFDQMGLI----DKKTMLEVERALHKI---------- 555

Qy    288 NDPKIV----------------KDDCGYKIDAQWVDDFHHAVH--AFITKEKDYYYQDFG 329
             || |:               | |      | : |:|  |:    |    | :    :|
Db    556 -DPTIILYGEPWGGWGAPIRFGKSDVAGTHVAAFNDEFRDAIRGSVFNPSVKGFVMGGYG 614

Qy    330 RIEDIEKTFKDVFVYDGKYSRYRGRTHGAPVGDLPPRKFVVFIQNHDQVGNRGNGERLSI 389
          :   |::       ||||  :             | | : : :   ||
Db    615 KETKIKRGVVGSINYDGKLIKSFA---------LDPEETINYAACHDN----------HT 655

Qy    390 LTDKTTYL--------------------MAATLYILSPYIPLIFMGEEYYETNPFFFFSD 429
          | ||  ||                    :|  : : |  :| :  |::: |   |  :
Db    656 LWDK-NYLAAKADKKKEWTEEELKNAQKLAGAILLTSQGVPFLHGGQDFCRTKN-FNDNS 713

Qy    430 FSDPVLIKGVREGRLKENNQMIDPQSEEAFLKSKLSWKIDEEVLDYYKQLINIRK----- 484
          :: |: | |             | : :  |:          :| :|:| || :||
Db    714 YNAPISING------------FDYERKLQFI----------DVFNYHKGLIKLRKEHPAF 751

Qy    485 RYNNCKRVKE 494
          |  | : :|:
Db    752 RLKNAEEIKK 761


RESULT   9
AMY2_DICTH
ID   AMY2_DICTH     STANDARD;      PRT;   562 AA.
AC   P14898;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   01-DEC-1992 (Rel. 24, Last annotation update)
DE   Alpha-amylase 2 (EC 3.2.1.1) (1,4-alpha-D-glucan glucanohydrolase).
GN   AMYB.
OS   Dictyoglomus thermophilum.
OC   Bacteria; Dictyoglomus group; Dictyoglomus.
OX   NCBI_TaxID=14;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 1-8.
RC   STRAIN=H-6-12;
RX   MEDLINE=88329076; PubMed=2458257;
RA   Horinouchi S., Fukusumi S., Ohshima T., Beppu T.;
RT   "Cloning and expression in Escherichia coli of two additional amylase
RT   genes of a strictly anaerobic thermophile, Dictyoglomus thermophilum,
RT   and their nucleotide sequences with extremely low
RT   guanine-plus-cytosine contents.";
RL   Eur. J. Biochem. 176:243-253(1988).
```

```
CC   -!- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
CC       linkages in oligosaccharides and polysaccharides.
CC   -!- PATHWAY: POLYSACCHARIDE DEGRADATION.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- MISCELLANEOUS: WHEN COMPARED TO AMYA, AMYB PRODUCED LARGER AMOUNTS
CC       OF REDUCING SUGAR.
CC   -!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC       KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X13199; CAA31586.1; -.
DR   PIR; S01312; S01312.
DR   HSSP; Q08751; 1BVZ.
DR   InterPro; IPR000461; Alpha_amylase.
DR   Pfam; PF00128; alpha-amylase; 1.
KW   Hydrolase; Glycosidase; Carbohydrate metabolism; Multigene family.
FT   ACT_SITE    338    338       BY SIMILARITY.
FT   ACT_SITE    404    404       BY SIMILARITY.
SQ   SEQUENCE   562 AA;  66999 MW;  9583342CFB7A369F CRC64;
```

```
  Query Match             10.7%;  Score 315;  DB 1;  Length 562;
  Best Local Similarity   21.7%;  Pred. No. 8.3e-15;
  Matches  128;  Conservative  105;  Mismatches  180;  Indels  178;  Gaps   24;

Qy       7 NIEKNKGIFKLWAPYVNSVKLKLSKKLIPMEKNDEGFFEVEIDDIEENLTYSYIIEDKRE 66
           ||:: |     |   : :   :  |: |:::|  | ::: : :: | :   |:  |  |
Db      72 NIQERK---TQWNDSIFNFSEIIKIKIPPLKEN--GLYQIHLYEMNEKIYEQYLSIDSFE 126

Qy      67 IPDPASRYQPLGVHDKSQLIRTDYQILDLGKVKIEDLIIYELHVGTFSQE---------- 116
            |                          :  |: ||| : :  |:::
Db     127 AP---------------------------LWSEESIIYHIFIDRFAKDEKEVEYSENL 157

Qy     117 -----GNFKGVIEKLDYLKDLGITGIELMPVAQFPGNRDWGYDGVFLYAVQNTYGGPWEL 171
                 || ||:: :|||:::|||  | : |:  |      |||    : :   :|   :|
Db     158 KEKLGGNLKGILSRLDYIENLGINTIWISPI--FKSTSYHGYDIEDYFEIDPIWGTKEDL 215

Qy     172 AKLVNEAHKRGIAVILDVVYNHIG----------PEGNYLLGLGPYFSDRYKTPWGL-- 218
           ||| ||  ||| :||| | ||:           : : |    :  : |:| :|:
Db     216 KKLVREAFNRGIRIILDFVPNHMSYKNPIFQKALKDKNSNLRSWFIFKGEDYETFFGVKS 275

Qy     219 --TFNFDDRGCDQVRKFILENVEYWFKTFKIDGLRLDAVHA---------IFDNSPK--- 264
             |  ::   :  :|:   :|| : | | | |:|  ||          ||  : |
Db     276 MPKINLKNK---EAIDYIINAAKYWIREFGISGYRMD--HATGPDINFWSIFYYNLKSEF 330

Qy     265 -------HILQEIAEKAHQLGKFVIAESDLNDPKIVKDDCGYKIDAQWVDDFHHAVHAFI 317
                 |::   |    :|||     |    ||::|   : |  :|
Db     331 PETFYFGEIVETPKETKKYVGKF-DGTLDFYLFKIIRD---FFIGKRW----------S 375

Qy     318 TKEKDYYYQDFGRIEDIEKTFKDVFVYDGKYSRYRGRTHGAPVGDLPPRKFVVFIQNHDQ 377
           |||       | :: |:|: |     |  |: |                  : |::|||
Db     376 TKE-------FVKMIDLEEKF-----YGNKFKR-----------------ISFLENHD- 404

Qy     378 VGNRGNGERLSILTDKTTYLMAATLYILSPYIPLIFMGEEYYETNPFFFFSDFSDPVLIK 437
               :   | :  ||    :|:        ||:|: |:|          | : |  :::
Db     405 -----SNRFLWVAKDKKLLRLASIFQFSINAIPIIYNGQE-------MGCSQYRD--ILE 450

Qy     438 GVREGRLKENNQMIDPQSEEAFLKSKLSWKIDEEVLDYYKQLINIRKRY---------- 486
           | |   | |: ::  | |::         | |:|::|:|:||: ||| :
Db     451 GNR--TLHEHARLPIPWSDD---------KQDKELIDFYRQLVKIRKSHPALYKGTFIPI 499

Qy     487 --NNCKRVKEVRREGNCITLIMEK----IGIIASFDDIVINSKITGNLLIG 531
             :    :|| : |   : : :|       :  :: |:   :  | :| :|
Db     500 FSDMISFIKETQEESILVLINIEDKEEIFNLNGTYRDLFSGNIYTNSLKLG 550
```

```
RESULT  10
GLGB_BACSU
ID   GLGB_BACSU     STANDARD;      PRT;   627 AA.
AC   P39118;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   1,4-alpha-glucan branching enzyme (EC 2.4.1.18) (Glycogen branching
DE   enzyme).
GN   GLGB.
OS   Bacillus subtilis.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Bacillus/Staphylococcus group; Bacillus.
OX   NCBI_TaxID=1423;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168;
RX   MEDLINE=94195107; PubMed=8145641;
RA   Kiel J.A.K.W., Boels J.M., Beldman G., Venema G.;
RT   "Glycogen in Bacillus subtilis: molecular characterization of an
RT   operon encoding enzymes involved in glycogen biosynthesis and
RT   degradation.";
RL   Mol. Microbiol. 11:203-218(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98048467; PubMed=9387221;
RA   Lapidus A., Galleron N., Sorokin A., Ehrlich S.D.;
RT   "Sequencing and functional annotation of the Bacillus subtilis genes
RT   in the 200 kb rrnB-dnaB region.";
RL   Microbiology 143:3431-3441(1997).
CC   -!- FUNCTION: CATALYSES THE FORMATION OF THE ALPHA-1,6-GLUCOSIDIC
CC       LINKAGES IN GLYCOGEN BY SCISSION OF A 1,4-ALPHA-LINKED
CC       OLIGOSACCHARIDE FROM GROWING ALPHA-1,4-GLUCAN CHAINS AND THE
CC       SUBSEQUENT TRANSFER OF THE OLIGOSACCHARIDE TO OTHER PARTS OF
CC       ALPHA-1,4-GLUCAN CHAINS.
CC   -!- CATALYTIC ACTIVITY: Formation of 1,6-glucosidic linkages of
CC       glycogen.
CC   -!- PATHWAY: THIRD STEP IN GLYCOGEN BIOSYNTHESIS.
CC   -!- INDUCTION: EXPRESSED EXCLUSIVELY ON MEDIA CONTAINING CARBON
CC       SOURCES THAT ALLOW EFFICIENT SPORULATION.
CC   -!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC       KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z25795; CAA81040.1; -.
DR   EMBL; AF008220; AAC00214.1; -.
DR   EMBL; Z99119; CAB15076.1; -.
DR   PIR; S36624; S36624.
DR   SubtiList; BG10907; glgB.
DR   InterPro; IPR000461; Alpha_amylase.
DR   InterPro; IPR004193; isoamylase_N.
DR   Pfam; PF00128; alpha-amylase; 1.
DR   Pfam; PF02922; isoamylase_N; 1.
KW   Glycogen biosynthesis; Transferase; Glycosyltransferase;
KW   Complete proteome.
FT   ACT_SITE    309    309       BY SIMILARITY.
FT   ACT_SITE    352    352       BY SIMILARITY.
FT   ACT_SITE    420    420       BY SIMILARITY.
SQ   SEQUENCE   627 AA;  73665 MW;  64B0A5553B6767BA CRC64;
```

```
  Query Match             10.6%;  Score 313.5;  DB 1;  Length 627;
  Best Local Similarity   23.4%;  Pred. No. 1.2e-14;
  Matches  143;  Conservative   98;  Mismatches  207;  Indels  163;  Gaps   30;

Qy       4 FGGNIEKNKG----IFKLWAPYVNSVKLK------LSKKLIPMEKNDEGFFEVEIDDIEE 53
```

```
Db      26 FGSHYRELNGKSGYEFCVWAPHASEVRVAGDFNSWSGEEHVMHRVNDNGIWTLFIPGIGE 85

Qy      54 NLTYSY-IIEDKREI---PDPASRYQPLGVH-----------------DKSQLIRTDYQ 91

Db      86 KERYKYEIVTNNGEIRLKADPYAIYSEVRPNTASLTYDLEGYSWQDQKWQKKQKAKTLYE 145

Qy      92 ILDLGKVKIEDLIIYELHVGTFSQEG-----NFKGVIEKL-DYLKDLGITGIELMPVAQF 145

Db     146 ---------KPVFIYELHLGSWKKHSDGRHYSYKELSQTLIPYIKKHGFTHIELLPVYEH 196

Qy     146 PGNRDWGYDGVFLYAVQNTYGGPWELAKLVNEAHKRGIAVILDVVYNHI--GPEGNYLLG 203

Db     197 PYDRSWGYQGTGYYSPTSRFGPPHDLMKFVDECHQQNIGVILDWVPGHFCKDAHGLYMFD 256

Qy     204 LGPYFS-----DRYKTPWGLTFNFDDRGCDQVRKFILENVEYWFKTFKIDGLRLDAVHAI 258

Db     257 GEPLYEYKEERDRENWLWG-TANF-DLGKPEVHSFLISNALYWAEFYHIDGFRVDAVANI 314

Qy     259 F---DNSPKH---ILQEIAEKAHQLGK------FVIAESDLNDPKI--VKDDCG----YK 300

Db     315 LYWPNQDERHTNPYAVDFLKKLNQTMREAYPHVMMIAEDSTEWPQVTGAVEEGGLGFHYK 374

Qy     301 IDAQWVDDFHHAVHAFITKEKDYYYQ--DFGRIEDIEKTFKDVFVYD----GKYSRYRGR 354

Db     375 WNMGWMNDVLKYMET-PPEERRHCHQLISFSLLYAFSEHFVLPFSHDEVVYGKKSL---- 429

Qy     355 THGAPVGDLPPRKFVVFIQNHDQVGNRGNGERLSILTDKTTYLMAATLYILSPYIPLIFM 414

Db     430 -----LNKMPGDYWQKFAQ----------------------YRLLLGYMTVHPGKKLIFM 462

Qy     415 GEEYYETNPFFFFSDFSDPVLIKGVREGRLKENNQMID--PQSEEA----------FLKS 462

Db     463 GSEFAQ------FDEWKD----------TEQLDWFLDSFPMHQKASVFTQDLLRFYQKS 505

Qy     463 KLSWKIDE-----EVLDYY--KQLINIRKRYNNCKRVKEVRREGNCITLIMEKIGIIASF 515

Db     506 KILYEHDHRAQSFEWIDVHNDEQSIFSFIRYG--------QKHG-------EALVIICNF 550

Qy     516 DDIVINSKITG 526

Db     551 TPVVYHQYDVG 561
```

RESULT 11
GLGB_STRAU
ID GLGB_STRAU STANDARD; PRT; 764 AA.
AC P52980;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE 1,4-alpha-glucan branching enzyme (EC 2.4.1.18) (Glycogen branching
DE enzyme).
GN GLGB.
OS Streptomyces aureofaciens.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1894;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CCM 3239 / ATCC 10762;
RX MEDLINE=94347823; PubMed=8068720;
RA Homerova D., Kormanec J.;
RT "Cloning of the putative glycogen branching enzyme gene, glgB, from
RT Streptomyces aureofaciens.";
RL Biochim. Biophys. Acta 1200:334-336(1994).
CC -!- CATALYTIC ACTIVITY: Formation of 1,6-glucosidic linkages of
CC glycogen.
CC -!- PATHWAY: THIRD STEP IN GLYCOGEN BIOSYNTHESIS.
CC -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC ---

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC ---
DR EMBL; L11647; AAA67437.1; -.
DR InterPro; IPR000461; Alpha_amylase.
DR InterPro; IPR001899; Gram_pos_anchor.
DR InterPro; IPR004193; isoamylase_N.
DR Pfam; PF00128; alpha-amylase; 1.
DR Pfam; PF02922; isoamylase_N; 1.
KW Glycogen biosynthesis; Transferase; Glycosyltransferase.
FT ACT_SITE 440 440 BY SIMILARITY.
FT ACT_SITE 493 493 BY SIMILARITY.
FT ACT_SITE 561 561 BY SIMILARITY.
SQ SEQUENCE 764 AA; 85325 MW; 6B45482E4A268ACF CRC64;

Query Match 10.0%; Score 296.5; DB 1; Length 764;
Best Local Similarity 24.6%; Pred. No. 2.5e-13;
Matches 125; Conservative 88; Mismatches 177; Indels 119; Gaps 27;

```
Qy       3 SFGGNIEKNKGI----FKLWAPYVNSVKLKLSKKL-------IPMEK-NDEGFFEVEIDD 50

Db     162 ALGSQPMEHQGVAGTRFTVWAP--NALGVRVTGDFSYWDAVAYPMRSLGASGVWELFLPG 219

Qy      51 IEENLTYSYII---EDKREI-PDPASRYQPLGVHDKSQLIRTDYQILDL------GKVKI 100

Db     220 VAEGALYKYEITRPDGGRTLRADPMARYAEVPPANASIVTASRYEWQDAEWMARRGALAP 279

Qy     101 ED--LIIYELHVGTFSQEGNFKGVIEKLD-YLKDLGITGIELMPVAQFPGNRDWGYDGVF 157

Db     280 HQAPMSVYELHLASWRPGLSYRQLAEQLPAYVKELGFTHVELMPVAEHPFGGSWGYQVTG 339

Qy     158 LYAVQNTYGGPWELAKLVNEAHKRGIAVILDVVYNHIGPEGNYLLGLGPYFSDR----YK 213

Db     340 FYAPTSRMGTPDDFRFLVDALHRAGIGVIVDWVPAHF-PRDDWALA---EFDGRPLYEHQ 395

Qy     214 TP-------WGLTFNFDDRGCDQVRKFILENVEYWFKTFKIDGLRLDAVHAIF------- 259

Db     396 DPRRAAHPDWG-TLEF-DYGRKEVRNFLVANAVYWCQEFHVDGLRADAVASMLYLDYSRD 453

Qy     260 --DNSPK-----------HILQEIAEKAHQL--GKFVIAESDLNDPKIVK--DDCGYKID 302

Db     454 EGDWSPNAHGGREDLDAVALLQEMNATVYRRFPGVVTIAEESTAWDGVTRPTDSGGLGFG 513

Qy     303 AQWVDDFHHAVHAFITKE---KDYYYQDFGRIEDIEKTFKDVFVYDGKYSRYRGRTHGAP 359

Db     514 LKWNMGWMHDTLRYVSKEPVHRKYHHHDM--------TFGMVYAFS-------------- 551

Qy     360 VGDLPPRKFVVFIQNHDQVGNRGNGERLSILT--------DKTTYLMAATLYILSPYIPL 411

Db     552 ------ENFVLPI-SHDEV---VHGKR-SLVSKMPGDWWQQRATHRAYLGFMWAHPGKQL 600

Qy     412 IFMGEEYYETN-------PFFFFSDFSDP 433

Db     601 LFMGQEFAQGSEWSETYGPDWWVLDSSYP 629
```

RESULT 12
GLGB_SYNY3
ID GLGB_SYNY3 STANDARD; PRT; 770 AA.
AC P52981;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 1,4-alpha-glucan branching enzyme (EC 2.4.1.18) (Glycogen branching
DE enzyme).
GN GLGB OR SLL0158.
OS Synechocystis sp. (strain PCC 6803).

```
OC   Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX   NCBI_TaxID=1148;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96127529; PubMed=8590279;
RA   Kaneko T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N.,
RA   Sugiura M., Tabata S.;
RT   "Sequence analysis of the genome of the unicellular cyanobacterium
RT   Synechocystis sp. strain PCC6803. I. Sequence features in the 1 Mb
RT   region from map positions 64% to 92% of the genome.";
RL   DNA Res. 2:153-166(1995).
CC   -!- CATALYTIC ACTIVITY: Formation of 1,6-glucosidic linkages of
CC       glycogen.
CC   -!- PATHWAY: THIRD STEP IN GLYCOGEN BIOSYNTHESIS.
CC   -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC       KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D63999; BAA10073.1; -.
DR   InterPro; IPR000461; Alpha_amylase.
DR   InterPro; IPR004193; isoamylase_N.
DR   Pfam; PF00128; alpha-amylase; 1.
DR   Pfam; PF02922; isoamylase_N; 2.
KW   Glycogen biosynthesis; Transferase; Glycosyltransferase;
KW   Complete proteome.
FT   ACT_SITE    433    433       BY SIMILARITY.
FT   ACT_SITE    486    486       BY SIMILARITY.
FT   ACT_SITE    554    554       BY SIMILARITY.
SQ   SEQUENCE   770 AA;  89527 MW;  A435AFCA7703FA8A CRC64;


  Query Match             9.5%;  Score 282;  DB 1;  Length 770;
  Best Local Similarity   28.9%;  Pred. No. 2.6e-12;
  Matches   89;  Conservative   46;  Mismatches   97;  Indels   76;  Gaps   14;

Qy      12 KGI-FKLWAPYVNSVKL-----KLSKKLIPMEKNDEGFFEVEIDDIEENLTYSYIIEDKR 65
           ||: | :|||   :| :        :|  | | :   :|: | ::    :| | |::
Db     140 KGVYFAVWAPNARNVSILGDFNNWDGRLHQMRKRNNMVWELFIPELGVGTSYKYEIKNWE 199

Qy      66 ----EIPDPASRYQPLGVHDKSQLIRTD---YQILDLGKVKIEDLI---------- 104
               |  ||   ||   |  |:  |  |   ||  |      || :
Db     200 GHIYEKTDPYGFYQE--VRPKTASIVADLDGYQWHD------EDWLEARRTSDPLSKPVS 251

Qy     105 IYELHVG-----------TFSQEG---------------NFKGVIEKL-DYLKDLGITG 136
           :||||:|           |   ||              :  :::||  |:|:|| |
Db     252 VYELHLGSWLHTAYDEPVKTLHGEGVPVEVSEWNTGARFLTYYELVDKLIPYVKELGYTH 311

Qy     137 IELMPVAQFPGNRDWGYDGVFLYAVQNTYGGPWELAKLVNEAHKRGIAVILDVVYNHIGP 196
           |||:|:|: | :  |||     ||  : :| | :    |:: |  || ||:| |  |
Db     312 IELLPIAEHPFDGSWGYQVTGYYAPTSRFGSPEDFMYFVDQCHLNGIGVIIDWVPGHFPK 371

Qy     197 EGN--------YLLGLGPYFSDRYKTPWG-LTFNFDDRGCDQVRKFILENVEYWFKTFKI 247
           :|:        :|   |       :|  || | ||:  | ::|| |:: |  :||  : |
Db     372 DGHGLAFFDGTHLYEHGDPRKGEHK-EWGTLIFNY---GRNEVRNFLVANALFWFDKYHI 427

Qy     248 DGLRLDAV 255
           ||:|:|||
Db     428 DGMRVDAV 435


RESULT  13
GLGX_HAEIN
ID   GLGX_HAEIN     STANDARD;      PRT;   659 AA.
AC   P45178;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Glycogen operon protein glgX (EC 3.2.1.-).
GN   GLGX OR HI1358.
OS   Haemophilus influenzae.
OC   Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC   Haemophilus.
OX   NCBI_TaxID=727;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=RD / KW20 / ATCC 51907;
RX   MEDLINE=95350630; PubMed=7542800;
RA   Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA   Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA   McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA   Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA   Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA   Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA   Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA   Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA   Venter J.C.;
RT   "Whole-genome random sequencing and assembly of Haemophilus
RT   influenzae Rd.";
RL   Science 269:496-512(1995).
CC   -!- FUNCTION: THIS PROTEIN MAY BE PART OF A GLYCOGEN BIOSYNTHETIC/
CC       CATABOLIC OPERON BUT IS NOT REQUIRED FOR GLYCOGEN SYNTHESIS
CC       (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC       KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U32815; AAC23005.1; -.
DR   HSSP; P10342; 1BF2.
DR   TIGR; HI1358; -.
DR   InterPro; IPR000461; Alpha_amylase.
DR   InterPro; IPR004193; isoamylase_N.
DR   Pfam; PF00128; alpha-amylase; 1.
DR   Pfam; PF02922; isoamylase_N; 1.
KW   Glycogen biosynthesis; Hydrolase; Glycosidase; Complete proteome.
FT   ACT_SITE    333    333       BY SIMILARITY.
FT   ACT_SITE    440    440       BY SIMILARITY.
SQ   SEQUENCE   659 AA;  75290 MW;  E2B71F610E8B1CAC CRC64;


  Query Match             9.4%;  Score 276.5;  DB 1;  Length 659;
  Best Local Similarity   24.2%;  Pred. No. 5.2e-12;
  Matches  114;  Conservative   70;  Mismatches  157;  Indels  131;  Gaps   22;

Qy     102 DLIIYELHVGTFSQ---------EGNFKGVIE--KLDYLKDLGITGIELMPV-------- 142
           : |:|||||  |||`         | : |:     | |||:||:| :||:||
Db     148 ETIVYELHVKGFSQLNEKIPAALRGTYTGLAHPVNLAYLKELGVTAVELLPVNFHINEPH 207

Qy     143 AQFPGNRD-WGYDGVFLYAVQNTYGGP----WELAKLVNEAHKRGIAVILDVVYNHIG-P 196
             |  | :: |||: : ::||:  |        |   :|   || || ||||||:||
Db     208 LQARGLQNYWGYNPLAMFAVEPKYAATNNPLAEFKTMVKAFHKAGIEVILDVVFNHSAES 267

Qy     197 EGNYLLGLGPYFSDR--------YKTPWGLTFNFDDRGC--------DQVRKFILENVEY 240
           |  |     | || |        ::   |   |:   ||        |  ||:::: : |
Db     268 EQTY-----PTFSQRGIDDQTYYWRNDQGRYINW--TGCGNMLNLSSDVGRKWVVDCLRY 320

Qy     241 WFKTFKIDGLRLDAVHAIFDNSPKHILQEIAEKAHQLGKFVIAESDLNDPKIVKD----- 295
           | :   ||| | |    :  ::|          : ||   :  |  | : |:: :
Db     321 WVEQCHIDGFRFDLATVLGRDTPDF------NSSAQLFTDIKNEPSLQNIKLIAEPWDIG 374

Qy     296 DCGYKID------AQWVDDFHHAVHAF---ITKEKDYYYQDFGRIEDIEK---------- 336
           ||::       |:| | |   :  |    : |   : : |    |: |
```

```
Db    375 HYGYQVGNFPSYFAEWNDRFRDDLCRFWLWKSGEIGAFAERFAGSSDLFKKNDRLPHTTL 434

Qy    337 ---------TFKDVFVYDGKYSRYRGRTHGAPVGDLPPRKFVVFIQNHDQVGNRGNGERL 387
                  | ||:  |: |::   |  :      :        :  ||   |  |: | |
Db    435 NFITAHDGFTLKDLVSYNQKHNETNGEENRDGRNE-------NYSYNH---GVEGSTESL 484

Qy    388 -----SILTDKTTYLMAATL--YILSPYIPLIFMGEEYYETNPFFFFSDFSDPVLIKGVR 440
                | : :  |:  : |   :|:   |::  |:|:  |
Db    485 SEPQKSAVENNRTFAQSGLLMSLLLANGTPMLLAGDEFGNTQ------------------ 526

Qy    441 EGRLKENNQMIDPQSEEAFLKSKLSW-KIDEEVLDYYKQLINIRKRYNNCKR 491
                ||      :|  :||    |   :||: :  || | :||:  :  :
Db    527 ----YGNNNAYCQDNEITWLK----WANFNEELFELTKQTIALRKQIGSLNK 570


RESULT  14
GLGB_MYCTU
ID   GLGB_MYCTU     STANDARD;      PRT;   731 AA.
AC   Q10625;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Probable 1,4-alpha-glucan branching enzyme (EC 2.4.1.18) (Glycogen
DE   branching enzyme).
GN   GLGB OR RV1326C OR MT1368 OR MTCY130.11C.
OS   Mycobacterium tuberculosis.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX   NCBI_TaxID=1773;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=H37RV;
RX   MEDLINE=98295987; PubMed=9634230;
RA   Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA   Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekaia F.,
RA   Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA   Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA   Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA   Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA   Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA   Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT   "Deciphering the biology of Mycobacterium tuberculosis from the
RT   complete genome sequence.";
RL   Nature 393:537-544(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CDC 1551 / Oshkosh;
RA   Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA   Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA   Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA   Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA   Bishai W.;
RT   "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT   laboratory strains.";
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- CATALYTIC ACTIVITY: Formation of 1,6-glucosidic linkages of
CC       glycogen.
CC   -!- PATHWAY: THIRD STEP IN GLYCOGEN BIOSYNTHESIS.
CC   -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC       KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; Z73902; CAA98090.1; -.
DR   EMBL; AE007010; AAK45632.1; -.
DR   TIGR; MT1368; -.
DR   TubercuList; Rv1326c; -.
DR   InterPro; IPR000461; Alpha_amylase.
DR   InterPro; IPR004193; isoamylase_N.
DR   Pfam; PF00128; alpha-amylase; 1.
DR   Pfam; PF02922; isoamylase_N; 1.
KW   Glycogen biosynthesis; Transferase; Glycosyltransferase;
KW   Complete proteome.
FT   ACT_SITE    411    411       BY SIMILARITY.
FT   ACT_SITE    464    464       BY SIMILARITY.
FT   ACT_SITE    532    532       BY SIMILARITY.
FT   CONFLICT    214    214       P -> A (IN REF. 2).
FT   CONFLICT    223    223       P -> Q (IN REF. 2).
SQ   SEQUENCE   731 AA;  81729 MW;  EE2BFEF765352617 CRC64;


  Query Match             9.2%;  Score 273;  DB 1;  Length 731;
  Best Local Similarity   24.7%;  Pred. No. 1e-11;
  Matches  118;  Conservative   71;  Mismatches  168;  Indels  120;  Gaps   23;

Qy     15 FKLWAPYVNSVKL-----KLSKKLIPME-KNDEGFFEVEIDDIEENLTYSY-------II 61
          | :|||    | |       :    ||      | :|:   |   :  | :       ::
Db    149 FAVWAPNAKGVSLIGEFNGWNGHEAPMRVLGPSGVWELFWPDFPCDGLYKFRVHGADGVV 208

Qy     62 EDKREIPDPASRYQPLGVHDKSQLIRTDYQILD--------LGKVKIEDLIIYELHVGTF 113
           |:    || :    :     |::  :||   |        |     | :  ||:|:|::
Db    209 TDR---ADPFAFGTEVPPQTASRVTSSDYTWGDDDWMAGRALRNPVNEAMSTYEVHLGSW 265

Qy    114 SQEGNFKGVIEKL-DYLKDLGITGIELMPVAQFPGNRDWGYDGVFLYAVQNTYGGPWELA 172
              ::: :  :| ||: | | | :||:|||: |    |||     ||  : :| | :
Db    266 RPGLSYRQLARELTDYIVDQGFTHVELLPVAEHPFAGSWGYQVTSYYAPTSRFGTPDDFR 325

Qy    173 KLVNEAHKRGIAVILDVVYNHIGPEGNYLLGLGPYFSDRYK-TP--------------WG 217
           ||:  |: || ||:| |  |  |:  : ||        |:  ||               ||
Db    326 ALVDALHQAGIGVIVDWVPAHF-PKDAWALG-------RFDGTPLYEHSDPKRGEQLDWG 377

Qy    218 LTFNFDDRGCDQVRKFILENVEYWFKTFKIDGLRLDA------------------VHAI 258
           |: | | |  :|| |:: |  || : | |||||:||                  ||
Db    378 -TYVF-DFGRPEVRNFLVANALYWLQEFHIDGLRVDAVASMLYLDYSRPEGGWTPNVHGG 435

Qy    259 FDN-SPKHILQEIAEKAHQL--GKFVIAE-----SDLNDPKIVKDDCGYKIDAQWVDDFH 310
           :|      |||:   ||::  |   |||     | :  |  :    |     :|   :
Db    436 RENLEAVQFLQEMNATAHKVAPGIVTIAEESTPWSGVTRPTNIG---GLGFSMKWNMGWM 492

Qy    311 HAVHAFITKE---KDYYYQDFGRIEDIEKTFKDVFVYDGKYSRYRGRTHGAPVGDLPPRK 367
          |    :::::   : |::          | ||  :: :   |
Db    493 HDTLDYVSRDPVYRSYHHH--------EMTFSMLYAFSENY------------------- 525

Qy    368 FVVFIQNHDQVGNRGNGERLSILTDKTTYLMAATLYILSPY------IPLIFMGEEY 418
            |   :||:| : | |      :     :: || |  |  |        |:|||:|:
Db    526 --VLPLSHDEVVH-GKGTLWGRMPG-NNHVKAAGLRSLLAYQWAHPGKQLLFMGQEF 578


RESULT  15
GLGB_HUMAN
ID   GLGB_HUMAN     STANDARD;      PRT;   702 AA.
AC   Q04446;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   1,4-alpha-glucan branching enzyme (EC 2.4.1.18) (Glycogen branching
DE   enzyme) (Brancher enzyme).
GN   GBE1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=93216700; PubMed=8463281;
RA   Thon V.J., Khalil M., Cannon J.F.;
RT   "Isolation of human glycogen branching enzyme cDNAs by screening
RT   complementation in yeast.";
```

```
RL   J. Biol. Chem. 268:7509-7513(1993).
CC   -!- FUNCTION: REQUIRED FOR SUFFICIENT GLYCOGEN ACCUMULATION. THE
CC       ALPHA 1-6 BRANCHES OF GLYCOGEN PLAY AN IMPORTANT ROLE IN
CC       INCREASING THE SOLUBILITY OF THE MOLECULE AND, CONSEQUENTLY, IN
CC       REDUCING THE OSMOTIC PRESSURE WITHIN CELLS.
CC   -!- CATALYTIC ACTIVITY: Formation of 1,6-glucosidic linkages of
CC       glycogen.
CC   -!- PATHWAY: THIRD STEP IN GLYCOGEN BIOSYNTHESIS.
CC   -!- SUBUNIT: MONOMER.
CC   -!- TISSUE SPECIFICITY: HIGHEST LEVELS FOUND IN LIVER AND MUSCLE.
CC   -!- DISEASE: DEFECTS IN GBE1 ARE THE CAUSE OF GLYCOGEN STORAGE DISEASE
CC       IV (GSD-IV) (ALSO KNOWN AS ANDERSEN'S DISEASE); A RARE FORM OF
CC       GLYCOGENOSIS CHARACTERIZED BY THE ACCUMULATION OF ABNORMALLY
CC       STRUCTURED GLYCOGEN THAT RESULTS IN EARLY ONSET HEPATIC CIRRHOSIS,
CC       CARDIAC ARREST AND NEUROMUSCULAR DISEASES. MOST CHILDREN WITH THIS
CC       CONDITION DIE BEFORE TWO YEARS OF AGE. NO TREATMENT APART FROM
CC       LIVER TRANSPLANTATION HAS BEEN FOUND TO PREVENT PROGRESSION OF THE
CC       DISEASE.
CC   -!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC       KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L07956; AAA58642.1; -.
DR   MIM; 232500; -.
DR   InterPro; IPR000461; Alpha_amylase.
DR   InterPro; IPR004193; isoamylase_N.
DR   Pfam; PF00128; alpha-amylase; 1.
DR   Pfam; PF02922; isoamylase_N; 1.
KW   Glycogen biosynthesis; Transferase; Glycosyltransferase;
KW   Glycogen storage disease.
FT   ACT_SITE    357    357       BY SIMILARITY.
FT   ACT_SITE    412    412       BY SIMILARITY.
FT   ACT_SITE    481    481       BY SIMILARITY.
SQ   SEQUENCE   702 AA;  80443 MW;  4DF3AA8D365A8FE3 CRC64;


  Query Match             9.1%;  Score 270;  DB 1;  Length 702;
  Best Local Similarity   21.9%;  Pred. No. 1.6e-11;
  Matches  147;  Conservative   90;  Mismatches  202;  Indels  232;  Gaps   29;

Qy       3 SFGGNIEKNKGIF-KLWAPYVNSVKLK-----LSKKLIPMEKNDEGFFEVEIDDIE---- 52
           ||| :   : |:: | |||    | |       :    | :| | | :|: |   :
Db      75 SFGVHRCADGGLYSKEWAPGAEGVFLTGDFNGWNPFSYPYKKLDYGKWELYIPPKQNKSV 134

Qy      53 ---------------------------------ENLTYSYIIEDKREIPDPASRYQPLG 78
                                              :|: | :|  |    |: :  ::
Db     135 LVPHGSKLKVVITSKSGEILYRISPWAKYVVREGDNVNYDWIHWD----PEHSYEFK--- 187

Qy      79 VHDKSQLIRTDYQILDLGKVKIEDLIIYELHVGTFSQEGNFKG----VIEKLDYLKDLGI 134
           | : :              |   | ||| |||  | ||            |  :| ||
Db     188 -HSRPK-------------KPRSLRIYESHVGISSHEGKVASYKHFTCNVLPRIKGLGY 232

Qy     135 TGIELMPVAQFPGNRDWGYDGVFLYAVQNTYGGPWELAKLVNEAHKRGIAVILDVV---- 190
             |:|| : :      :||     :|  : || | || :||: ||  || |:||||
Db     233 NCIQLMAIMEHAYYASFGYQITSFFAASSRYGTPEELQELVDTAHSMGIIVLLDVVHSHA 292

Qy     191 -----------------YNHIGPEGNYLLGLGPYFSDRYKTPWGLTFNFDDRGCDQVRKF 233
                            | | || | : |     |:    : |            :| :|
Db     293 SKNSADGLNMFDGTDSCYFHSGPRGTHDLWDSRLFA---YSSW-----------EVLRF 337

Qy     234 ILENVEYWFKTFKIDGLRLDAVHAIFDNSPKHILQEIAEKAHQLGK-FVIAESDLNDPKI 292
           :| |: :| : :: || | | | ::  :             | :|: |    |:    ::
Db     338 LLSNIRWWLEEYRFDGFRFDGVTSMLYHH-----------HGVGQGFSGDYSEYFGLQV 385

Qy     293 VKDDCGYKIDAQWVDDFHHAVHAF----ITKEKD----------------YYYQDFGRI 331
           :|   | : |      :| ||      ||  :|                 : |:    |

Db     386 DEDALTYLMLA------NHLVHTLCPDSITIAEDVSGMPALCSPISQGGGGFDYRLAMAI 439

Qy     332 ED----IEKTFKDV------FVYDGKYSRYRGRTHGAPVGDLPPRKFVVFIQNHDQVGNR 381
             |    : | |||        ||     ||                | : : ::|||
Db     440 PDKWIQLLKEFKDEDWNMGDIVYTLTNRRYL------------EKCIAYAESHDQ---- 482

Qy     382 GNGERLSILTDKTT--YLMAATLY----ILSPYIPLI---------------FMGEEYY 419
                  ::: ||:   :|| | :|    :|:|: |:|                  || |
Db     483 ------ALVGDKSLAFWLMDAEMYTNMSVLTPFTPVIDRGIQLHKMIRLITHGLGGEGYL 536

Qy     420 ETNPFFFFSDFSDPVLIKGVREGRLKENNQMIDPQSEEAFLKSKLSWKIDEEVLDYYKQL 479
                 |  ::|  |  :   |:|    ||        |::  ::  : : :: |  || |
Db     537 N----FMGNEFGHPEWLDFPRKG----NN--------ESYHYARRQFHLTDDDLLRYKFL 580

Qy     480 INIRKRYNNCKR----------VKEVRREGNCITLIMEKIGIIASFDDIVINSKITGNL 528
             |  :  |  :            | | : ||| | :  |: |::  |:     ||   :
Db     581 NNFDRDMNRLEERYGWLAAPQAYVSE-KHEGNKI-IAFERAGLLFIFN--FHPSKSYTDY 636

Qy     529 LIGIGFPKKLK 539
           :|   | | |
Db     637 RVGTALPGKFK 647


Search completed: July 15, 2002, 12:24:34
Job time: 386 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: July 15, 2002, 12:24:08 ; Search time 55.94 Seconds
(without alignments)
1719.435 Million cell updates/sec

Title: US-09-298-924-8
Perfect score: 2954
Sequence: 1 MFSFGGNIEKNKGIFKLWAP.....KLKKDELIKVNRGVGVYQLE 556

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_19:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2954	100.0	556	1	Q53641	Q53641 sulfolobus
2	1784.5	60.4	561	17	P95867	P95867 sulfolobus
3	1748.5	59.2	559	1	Q55088	Q55088 sulfolobus
4	1711.5	57.9	559	1	Q9UWN9	Q9uwn9 sulfolobus
5	1290	43.7	359	17	Q973H4	Q973h4 sulfolobus
6	942.5	31.9	581	2	Q9ADI5	Q9adi5 streptomyce
7	865.5	29.3	583	16	Q9I1V1	Q9i1v1 pseudomonas
8	835.5	28.3	600	16	Q9RX51	Q9rx51 deinococcus
9	792.5	26.8	601	16	Q92U63	Q92u63 rhizobium m
10	645	21.8	217	17	Q973H5	Q973h5 sulfolobus
11	493	16.7	552	2	Q44528	Q44528 anabaena va
12	420	14.2	440	2	Q93Q35	Q93q35 myxococcus
13	360	12.2	652	2	Q59242	Q59242 bacillus st
14	333.5	11.3	707	16	P73608	P73608 synechocyst
15	332.5	11.3	1280	16	Q97SQ7	Q97sq7 streptococc
16	331	11.2	1142	2	Q93UZ9	Q93uz9 bacillus sp
17	326.5	11.1	1287	2	Q9F930	Q9f930 streptococc
18	326	11.0	718	16	O34587	O34587 bacillus su
19	322.5	10.9	422	2	Q59243	Q59243 bacillus st
20	319.5	10.8	718	17	P95868	P95868 sulfolobus
21	311	10.5	718	2	O69008	O69008 thermus sp.
22	310.5	10.5	642	16	Q97QS8	Q97qs8 streptococc
23	306.5	10.4	666	16	O84046	O84046 chlamydia t
24	306	10.4	825	2	Q59319	Q59319 caldocellum
25	303.5	10.3	664	16	Q9Z8F5	Q9z8f5 chlamydia p
26	302.5	10.2	1165	16	Q99XX8	Q99xx8 streptococc
27	302	10.2	783	10	O04196	O04196 arabidopsis
28	298.5	10.1	621	2	Q93HU3	Q93hu3 rhodothermu
29	298.5	10.1	716	17	Q973H3	Q973h3 sulfolobus
30	298	10.1	713	1	O05152	O05152 sulfolobus
31	296.5	10.0	818	10	Q41742	Q41742 zea mays (m
32	296	10.0	630	16	O66936	O66936 aquifex aeo
33	295	10.0	666	16	Q9PKZ6	Q9pkz6 chlamydia m
34	295	10.0	789	10	O22637	O22637 zea mays (m
35	293.5	9.9	717	16	Q9K7U5	Q9k7u5 bacillus ha
36	292.5	9.9	660	1	Q9HHC8	Q9hhc8 thermococcu
37	292.5	9.9	741	3	Q9P5P3	Q9p5p3 neurospora
38	292	9.9	1072	16	Q9K6N1	Q9k6n1 bacillus ha
39	289	9.8	720	16	Q97FP9	Q97fp9 clostridium
40	289	9.8	741	2	Q59832	Q59832 streptomyce
41	288	9.7	741	2	Q9KY06	Q9ky06 streptomyce
42	286.5	9.7	666	16	Q9KNE8	Q9kne8 vibrio chol
43	286.5	9.7	668	2	P71095	P71095 bacteroides
44	284	9.6	733	10	O80403	O80403 oryza sativ
45	284	9.6	1938	2	P70983	P70983 bacillus sp

ALIGNMENTS

RESULT 1
Q53641
ID Q53641 PRELIMINARY; PRT; 556 AA.
AC Q53641; O08279; O08064;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE ALPHA-AMYLASE (FRAGMENT).
GN TREZ.
OS Sulfolobus acidocaldarius.
OC Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.
OX NCBI_TaxID=2285;
RN [1]
RP SEQUENCE FROM N.A.
RA Kato M., Kettoku M., Miura Y., Komeda T., Konishi Y., Shindo K.,
RA Kobayashi K., Iwamatsu A.;
RL Submitted (SEP-1995) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC33909;
RX MEDLINE=97135071; PubMed=8980629;
RA Maruta K., Mitsuzumi H., Nakada T., Kubota M., Chaen H., Fukuda S.,
RA Sugimoto T., Kurimoto M.;
RT "Cloning and sequencing of a cluster of genes encoding novel enzymes
RT of trehalose biosynthesis from thermophilic archaebacterium Sulfolobus
RT acidocaldarius.";
RL Biochim. Biophys. Acta 1291:177-181(1996).
DR EMBL; D64131; BAA11011.1; -.
DR EMBL; D83245; BAA11863.1; -.
DR InterPro; IPR000461; Alpha_amylase.
DR InterPro; IPR004193; isoamylase_N.
DR InterPro; IPR003662; sub_transporter.
DR Pfam; PF00128; alpha-amylase; 1.
DR Pfam; PF02922; isoamylase_N; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; UNKNOWN_1.
KW Hydrolase.
FT NON_TER 556 556
SQ SEQUENCE 556 AA; 64373 MW; 4215B45E6C8ED4E7 CRC64;

```
  Query Match           100.0%;  Score 2954;  DB 1;  Length 556;
  Best Local Similarity  100.0%;  Pred. No. 4.7e-189;
  Matches  556;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MFSFGGNIEKNKGIFKLWAPYVNSVKLKLSKKLIPMEKNDEGFFEVEIDDIEENLTYSYI 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MFSFGGNIEKNKGIFKLWAPYVNSVKLKLSKKLIPMEKNDEGFFEVEIDDIEENLTYSYI 60

Qy      61 IEDKREIPDPASRYQPLGVHDKSQLIRTDYQILDLGKVKIEDLIIYELHVGTFSQEGNFK 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 IEDKREIPDPASRYQPLGVHDKSQLIRTDYQILDLGKVKIEDLIIYELHVGTFSQEGNFK 120

Qy     121 GVIEKLDYLKDLGITGIELMPVAQFPGNRDWGYDGVFLYAVQNTYGGPWELAKLVNEAHK 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 GVIEKLDYLKDLGITGIELMPVAQFPGNRDWGYDGVFLYAVQNTYGGPWELAKLVNEAHK 180

Qy     181 RGIAVILDVVYNHIGPEGNYLLGLGPYFSDRYKTPWGLTFNFDDRGCDQVRKFILENVEY 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 RGIAVILDVVYNHIGPEGNYLLGLGPYFSDRYKTPWGLTFNFDDRGCDQVRKFILENVEY 240

Qy     241 WFKTFKIDGLRLDAVHAIFDNSPKHILQEIAEKAHQLGKFVIAESDLNDPKIVKDDCGYK 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 WFKTFKIDGLRLDAVHAIFDNSPKHILQEIAEKAHQLGKFVIAESDLNDPKIVKDDCGYK 300

Qy     301 IDAQWVDDFHHAVHAFITKEKDYYYQDFGRIEDIEKTFKDVFVYDGKYSRYRGRTHGAPV 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 IDAQWVDDFHHAVHAFITKEKDYYYQDFGRIEDIEKTFKDVFVYDGKYSRYRGRTHGAPV 360

Qy     361 GDLPPRKFVVFIQNHDQVGNRGNGERLSILTDKTTYLMAATLYILSPYIPLIFMGEEYYE 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 GDLPPRKFVVFIQNHDQVGNRGNGERLSILTDKTTYLMAATLYILSPYIPLIFMGEEYYE 420

Qy     421 TNPFFFFSDFSDPVLIKGVREGRLKENNQMIDPQSEEAFLKSKLSWKIDEEVLDYYKQLI 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 TNPFFFFSDFSDPVLIKGVREGRLKENNQMIDPQSEEAFLKSKLSWKIDEEVLDYYKQLI 480

Qy     481 NIRKRYNNCKRVKEVRREGNCITLIMEKIGIIASFDDIVINSKITGNLLIGIGFPKKLKK 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 NIRKRYNNCKRVKEVRREGNCITLIMEKIGIIASFDDIVINSKITGNLLIGIGFPKKLKK 540

Qy     541 DELIKVNRGVGVYQLE 556
           ||||||||||||||||
Db     541 DELIKVNRGVGVYQLE 556


RESULT   2
P95867
ID   P95867      PRELIMINARY;      PRT;   561 AA.
AC   P95867;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   ALPHA-AMYLASE PRECURSOR (MALTO-OLIGOSYLTREHALOSE TREHALOHYDROLASE)
DE   (TREZ) (EC 3.2.1.141).
GN   TREZ.
OS   Sulfolobus solfataricus.
OC   Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.
OX   NCBI_TaxID=2287;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 35092 / DSM 1617 / P2;
RX   MEDLINE=21332296; PubMed=11427726;
RA   She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
RA   Awayez M.J., Chan-Weiher C.C.-Y., Clausen I.G., Curtis B.A.,
RA   De Moors A., Erauso G., Fletcher C., Gordon P.M.K.,
RA   Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
RA   Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
RA   Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
RA   Garrett R.A., Ragan M.A., Sensen C.W., Van der Oost J.;
RT   "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
DR   EMBL; Y08256; CAA69503.1; -.
DR   EMBL; AE006815; AAK42272.1; -.
DR   InterPro; IPR000461; Alpha_amylase.
DR   InterPro; IPR004193; isoamylase_N.
DR   Pfam; PF00128; alpha-amylase; 1.
DR   Pfam; PF02922; isoamylase_N; 1.
KW   Hydrolase; Glycosidase; Complete proteome.
SQ   SEQUENCE   561 AA;  64370 MW;  B00E403020F6B242 CRC64;


  Query Match           60.4%;  Score 1784.5;  DB 17;  Length 561;
  Best Local Similarity  59.2%;  Pred. No. 4.9e-111;
  Matches  332;  Conservative   90;  Mismatches  130;  Indels    9;  Gaps    6;

Qy       3 SFGGNIEKNKGIFKLWAPYVNSVKLK-LSKKLIPMEKNDEGFFEVEIDDIEENLTYSYII 61
             :||  ::::   | |||||   |||| |:: :  ||::|:|:| : :|::     | ||:
Db       2 TFGYKLDEDGVTFNLWAPYQRKVKLKILNRGIYEMERDDKGYFTITLDNVRVGDRYKYIL 61

Qy      62 EDKREIPDPASRYQPLGVHDKSQLIRTDYQILDLGKVKI--EDLIIYELHVGTFSQEGNF 119
           :|  |:|||||||| |||  |::|  |::  |   ||:  |||:|||||:|||: || |
Db      62 DDNSEVPDPASRYQPEGVHGYSEIISPDFEWDDENSVKVKREDLVIYELHIGTFTSEGTF 121

Qy     120 KGVIEKLDYLKDLGITGIELMPVAQFPGNRDWGYDGVFLYAVQNTYGGPWELAKLVNEAH 179
           :|||:||:|||:||:| ||:||:||||| :|||||||:||||||:||||    |||||||
Db     122 EGVIKKLNYLKELGVTAIEIMPIAQFPGKKDWGYDGVYLYAVQNSYGGPSGFRKLVNEAH 181

Qy     180 KRGIAVILDVVYNHIGPEGNYLLGLGPYFSDRYKTPWGLTFNFDDRGCDQVRKFILENVE 239
           | |:||||||||:|||||:: |||||::|||||||||||||| | |:||||||||||
Db     182 KLGLAVILDVVYNHVGPEGNYMVKLGPYFSEKYKTPWGLTFNFDDAGSDEVRKFILENVE 241

Qy     240 YWFKTFKIDGLRLDAVHAIFDNSPKHILQEIAEKAHQLGKFVIAESDLNDPKIV--KDDC 297
           ||   | :|| |||||||| ||||||||::||:  |:  | ||||||||||::|  |: |
Db     242 YWINEFHVDGFRLDAVHAIIDNSPKHILEDIADVVHKYDKIVIAESDLNDPRVVNPKEKC 301

Qy     298 GYKIDAQWVDDFHHAVHAFITKEKDYYYQDFGRIEDIEKTFKDVFVYDGKYSRYRGRTHG 357
           || ||||||||||||:|||:| |:  || ||| | || |::||||:|||||| :| :|||
Db     302 GYNIDAQWVDDFHHAIHAFLTGERQGYYSDFGSIGDIVKSYKDVFIYDGKYSNFRRKTHG 361

Qy     358 APVGDLPPRKFVVFIQNHDQVGNRGNGERLSILTDKTTYLMAATLYILSPYIPLIFMGEE 417
             ||||   ||||:|||||||||||| ||||  | || :| :|| |||||||||:||||||
Db     362 KSVGDLDGCKFVVYIQNHDQVGNRGGGERLIKLVDKESYKIAAALYILSPYIPMIFMGEE 421

Qy     418 YYETNPFFFFSDFSDPVLIKGVREGRLKENNQMIDPQSEEAFLKSKLSWKIDEEVLDYYK 477
           | | |||::|||||||  ||:|||||| :|| |  ||||:  |  |||||||::::| :||
Db     422 YGEENPFYYFSDFSDPKLIQGVREGRRRENGQETDPQSDCTFNDSKLSWKINDDILSFYK 481

Qy     478 QLINIRKRYN-NCKRVKEVRREGNCITLIMEKIGIIASFDDIVINSKITGNLLI--GIGF 534
           || ||| |   | |   |      :|:       :  |   ||  | :| |::     |
Db     482 SLIKIRKEYGLACNRKLSVENGNYWLTVKGNGCLAVYVFSKSVIEMKYSGTLVLSSNSSF 541

Qy     535 PKKLKKDELIKVNRGVGVYQL 555
           | :: : :  ::::|  :|:|
Db     542 PSQITESK-YELDKGFALYKL 561


RESULT   3
Q55088
ID   Q55088      PRELIMINARY;      PRT;   559 AA.
AC   Q55088;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   ALPHA-AMYLASE.
OS   Sulfolobus solfataricus.
OC   Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.
OX   NCBI_TaxID=2287;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=KM1;
RA   Kato M., Kettoku M., Miura Y., Komeda T., Konishi Y., Shindo K.,
RA   Kobayashi K., Iwamatsu A.;
RT   "The gene analysis of the new amylases from the hyper thermophilic
RT   archae Sulfolobus.";
RL   Submitted (SEP-1995) to the EMBL/GenBank/DDBJ databases.
```

```
DR   EMBL; D64130; BAA11010.1; -.
DR   InterPro; IPR000461; Alpha_amylase.
DR   InterPro; IPR004193; isoamylase_N.
DR   Pfam; PF00128; alpha-amylase; 1.
DR   Pfam; PF02922; isoamylase_N; 1.
SQ   SEQUENCE   559 AA;  64790 MW;  79FBE23A7CD38B4E CRC64;


  Query Match           59.2%;  Score 1748.5;  DB 1;  Length 559;
  Best Local Similarity   58.9%;  Pred. No. 1.2e-108;
  Matches  329;  Conservative   87;  Mismatches  136;  Indels    7;  Gaps    5;

Qy       3 SFGGNIEKNKGIFKLWAPYVNSVKLK-LSKKLIPMEKNDEGFFEVEIDDIEENLTYSYII 61
           :|   |: |: || |||||  ||||| | | |  ||::::|:| : :::::    | |::
Db       2 TFAYKIDGNEVIFTLWAPYQKSVKLKVLEKGLYEMERDEKGYFTITLNNVKVRDRYKYVL 61

Qy      62 EDKREIPDPASRYQPLGVHDKSQLIRTDYQILDLGKVKIEDLIIYELHVGTFSQEGNFKG 121
           :|  ||||||||||| ||| ||:|:   :  :   :| |||||||:|||||: || |:|
Db      62 DDASEIPDPASRYQPEGVHGPSQIIQESKEFNNETFLKKEDLIIYEIHVGTFTPEGTFEG 121

Qy     122 VIEKLDYLKDLGITGIELMPVAQFPGNRDWGYDGVFLYAVQNTYGGPWELAKLVNEAHKR 181
           || |||||||||| ||:||:|||| |||||||:||||||:||||    |||:|||| :
Db     122 VIRKLDYLKDLGITAIEIMPIAQFPGKRDWGYDGVYLYAVQNSYGGPEGFRKLVDEAHKK 181

Qy     182 GIAVILDVVYNHIGPEGNYLLGLGPYFSDRYKTPWGLTFNFDDRGCDQVRKFILENVEYW 241
           |: |||||||||:||||||:: |||||| :|||||||||||    |:||||||||||
Db     182 GLGVILDVVYNHVGPEGNYMVKLGPYFSQKYKTPWGLTFNFDDAESDEVRKFILENVEYW 241

Qy     242 FKTFKIDGLRLDAVHAIFDNSPKHILQEIAEKAHQLGKFVIAESDLNDPKIV--KDDCGY 299
           | : :|| |||||||| | ||||||:|||:  |:  : ||||||||||::|  |: |||
Db     242 IKEYNVDGFRLDAVHAIIDTSPKHILEEIADVVHKYNRIVIAESDLNDPRVVNPKEKCGY 301

Qy     300 KIDAQWVDDFHHAVHAFITKEKDYYYQDFGRIEDIEKTFKDVFVYDGKYSRYRGRTHGAP 359
           |||||||||||::||::| |:  || ||| ::|| |::|||||||||| :| :||| |
Db     302 NIDAQWVDDFHHSIHAYLTGERQGYYTDFGNLDDIVKSYKDVFVYDGKYSNFRRKTHGEP 361

Qy     360 VGDLPPRKFVVFIQNHDQVGNRGNGERLSILTDKTTYLMAATLYILSPYIPLIFMGEEYY 419
           ||:|    |||:|||||||||| |||:  | |: :| :|| ||:|||||:|||||||
Db     362 VGELDGCNFVVYIQNHDQVGNRGKGERIIKLVDRESYKIAAALYLLSPYIPMIFMGEEYG 421

Qy     420 ETNPFFFFSDFSDPVLIKGVREGRLKENNQMIDPQSEEAFLKSKLSWKIDEEVLDYYKQL 479
           | |||:|||||||  ||:|||||| ||| |  ||| |  |  |||||||||:  :|| |
Db     422 EENPFYFFSDFSDSKLIQGVREGRKKENGQDTDPQDESTFNASKLSWKIDEEIFSFYKIL 481

Qy     480 INIRKRYN-NCKRVKEVRREGNCITLIMEKIGIIASFDDIVINSKITGNLLIGI--GFPK 536
           | :||  :  | |   |    | : :  :  :  |    |  | :| ||:     ||:
Db     482 IKMRKELSIACDRRVNVVNGENWLIIKGREYFSLYVFSKSSIEVKYSGTLLLSSNNSFPQ 541

Qy     537 KLKKDELIKVNRGVGVYQL 555
            ::: :  : ::|  :|:|
Db     542 HIEEGK-YEFDKGFALYKL 559


RESULT   4
Q9UWN9
ID   Q9UWN9      PRELIMINARY;      PRT;   559 AA.
AC   Q9UWN9;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   MALTOOLIGOSYL TREHALOSE TREHALOHYDROLASE.
GN   TREZ.
OS   Sulfolobus shibatae.
OC   Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.
OX   NCBI_TaxID=2286;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Weiting Y., Wei C., Hui W., Li L., Cheng J.;
RT   "Maltooligosyl trehalose trehalohydrolase from Sulfolobus shibatae.";
RL   Submitted (NOV-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF201335; AAF17553.1; -.
DR   InterPro; IPR000461; Alpha_amylase.
DR   InterPro; IPR004193; isoamylase_N.
DR   Pfam; PF00128; alpha-amylase; 1.
DR   Pfam; PF02922; isoamylase_N; 1.
KW   Hydrolase.
SQ   SEQUENCE   559 AA;  64619 MW;  8C0F0F38A5AE2846 CRC64;


  Query Match           57.9%;  Score 1711.5;  DB 1;  Length 559;
  Best Local Similarity   58.0%;  Pred. No. 3.6e-106;
  Matches  324;  Conservative   87;  Mismatches  141;  Indels    7;  Gaps    5;

Qy       3 SFGGNIEKNKGIFKLWAPYVNSVKLK-LSKKLIPMEKNDEGFFEVEIDDIEENLTYSYII 61
           :||  |: |: || |||||  ||||| | | |  ||::|:|:| : :::::    | |::
Db       2 TFGYKIDGNEVIFTLWAPYQKSVKLKVLEKGLYEMERDDKGYFTITLNNVKVGDRYKYVL 61

Qy      62 EDKREIPDPASRYQPLGVHDKSQLIRTDYQILDLGKVKIEDLIIYELHVGTFSQEGNFKG 121
           :|  ||||||||||| |||  ||:|:   :  :   :| |||||||:|||||: || |:|
Db      62 DDASEIPDPASRYQPEGVHGPSQIIQESKEFNNETFLKKEDLIIYEIHVGTFAPEGTFEG 121

Qy     122 VIEKLDYLKDLGITGIELMPVAQFPGNRDWGYDGVFLYAVQNTYGGPWELAKLVNEAHKR 181
           || |||||||||| ||:||:|||| |||||||:||||||:||||    |||:|||| :
Db     122 VIRKLDYLKDLGITAIEIMPIAQFPGKRDWGYDGVYLYAVQNSYGGPEGFRKLVDEAHMK 181

Qy     182 GIAVILDVVYNHIGPEGNYLLGLGPYFSDRYKTPWGLTFNFDDAESDEVRKFILENVEYW 241
           |: |||||||||:||||||:: |||||| :|||||||||||    |:||||||||||
Db     182 GLGVILDVVYNHVGPEGNYMVKLGPYFSQKYKTPWGLTFNFDDAESDEVRKFILENVEYW 241

Qy     242 FKTFKIDGLRLDAVHAIFDNSPKHILQEIAEKAHQLGKFVIAESDLNDPKIV--KDDCGY 299
           | : :|| |||||:| | ||||||:|||:  |:  : ||||||||||::|  |: |||
Db     242 IKEYNVDGFRLDAVHSIIDTSPKHILEEIADVVHKYNRIVIAESDLNDPRVVNPKEKCGY 301

Qy     300 KIDAQWVDDFHHAVHAFITKEKDYYYQDFGRIEDIEKTFKDVFVYDGKYSRYRGRTHGAP 359
           |||||||||||::||::| |:  || ||| ::|| |::|||||||||| :| :||| |
Db     302 NIDAQWVDDFHHSIHAYLTGERQGYYLDFGNLDNIVKSYKDVFVYDGKYSKFRRKTHGKP 361

Qy     360 VGDLPPRKFVVFIQNHDQVGNRGNGERLSILTDKTTYLMAATLYILSPYIPLIFMGEEYY 419
           ||:|    |||:|||||||||| |||:  | || :|  || |||||  | :|||||||
Db     362 VGELDGCNFVVYIQNHDQVGNRGKGERIIKLVDKESYKNAAALYILSSIISMIFMGEEYG 421

Qy     420 ETNPFFFFSDFSDPVLIKGVREGRLKENNQMIDPQSEEAFLKSKLSWKIDEEVLDYYKQL 479
           |  ||:|||||||  ||:|:|||| ||| |  ||| |  |  |||||||||:  :|| |
Db     422 EQKPFYFFSDFSDTKLIQGLREGRKKENGQDTDPQDESTFNASKLSWKIDEEIFSFYKIL 481

Qy     480 INIRKRYN-NCKRVKEVRREGNCITLIMEKIGIIASFDDIVINSKITGNLLIGI--GFPK 536
           | :|   :  | |    |    | : :  :  :  |    |  | :| ||:     ||:
Db     482 IKMRNELSIACDRRVYVENGENWLIIKGREYLSLCVFSRSSIEVKYSGTLLLSSNNSFPQ 541

Qy     537 KLKKDELIKVNRGVGVYQL 555
            ::: :  : ::|  :|:|
Db     542 HIEEGQ-YEFDKGFALYKL 559


RESULT   5
Q973H4
ID   Q973H4      PRELIMINARY;      PRT;   359 AA.
AC   Q973H4;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PUTATIVE ALPHA-AMYLASE.
GN   ST0927.
OS   Sulfolobus tokodaii.
OC   Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.
OX   NCBI_TaxID=111955;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JCM 10545 / 7;
RX   PubMed=11572479;
RA   Kawarabayasi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
RA   Sekine M., Baba S.-I., Ankai A., Kosugi H., Hosoyama A., Fukui S.,
RA   Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
RA   Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,
RA   Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,
RA   Oshima T., Kikuchi H.;
```

```
RT   "Complete genome sequence of an aerobic thermoacidophilic
RT   Crenarchaeon, Sulfolobus tokodaii strain7.";
RL   DNA Res. 8:123-140(2001).
DR   EMBL; AP000984; BAB65939.1; -.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   359 AA;  41600 MW;  4D9F429099E81665 CRC64;


  Query Match             43.7%;  Score 1290;  DB 17;  Length 359;
  Best Local Similarity   66.0%;  Pred. No. 2.6e-78;
  Matches  235;  Conservative   53;  Mismatches   66;  Indels    2;  Gaps    2;

Qy    201 LLGLGPYFSDRYKTPWGLTFNFDDRGCDQVRKFILENVEYWFKTFKIDGLRLDAVHAIFD 260
Db      1 MMFLGPYFSQKYKTPWGLTFNFDDAYSDEVRKFILDNVRYWFNVFHVDGLRLDAVHAIYD 60

Qy    261 NSPKHILQEIAEKAHQLGKFVIAESDLNDPKIVKDDCGYKIDAQWVDDFHHAVHAFITKE 320
Db     61 SSPKHILQDIAELSHSLGKFVIAESDLNDPKIIDDKCGYKIDAQWVDDFHHSVHAYLTGE 120

Qy    321 KDYYYQDFGRIEDIEKTFKDVFVYDGKYSRYRGRTHGAPVGDLPPRKFVVFIQNHDQVGN 380
Db    121 RNSYYSDFGTLDDIVKAFKDVFVYDGKYSNFRKKTHGAPVGNLNACKFVVYIQNHDQVGN 180

Qy    381 RGNGERLSILTDKTTYLMAATLYILSPYIPLIFMGEEYYETNPFFFFSDFSDPVLIKGVR 440
Db    181 RGNGDRLSTLVDKQSYMIASALYLLSPFIPMIFMGEEYYEKNPFLFFSDFSDPNLIKGVR 240

Qy    441 EGRLKENNQMIDPQSEEAFLKSKLSWKIDEEVLDYYKQLINIRKRY-NNCKRVKEVRREG 499
Db    241 EGRLRDNGQSIDPQSDEAFLKSKLSWNIDKEMLEYYRTLIKVRKEFGNKCSRRVEVDKTD 300

Qy    500 NCITLIMEKIGIIASFDDIVINSKITGNLLIGIG-FPKKLKKDELIKVNRGVGVYQ 554
Db    301 NWLVIGLEKIFLLVAFSDTGLILRYNGELLFSTSDFPKRIVSSKEIKVKKGVGVYK 356


RESULT   6
Q9ADI5
ID   Q9ADI5      PRELIMINARY;      PRT;   581 AA.
AC   Q9ADI5;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE   PUTATIVE ALPHA AMYLASE.
GN   SCBAC1A6.02C.
OS   Streptomyces coelicolor.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=1902;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3(2);
RA   Seeger K.J., Harris D.;
RL   Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3(2);
RA   Cerdeno A.M., Parkhill J., Barrell B.G., Rajandream M.A.;
RL   Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3(2);
RX   MEDLINE=97000351; PubMed=8843436;
RA   Redenbach M., Kieser H.M., Denapaite D., Eichner A., Cullum J.,
RA   Kinashi H., Hopwood D.A.;
RT   "A set of ordered cosmids and a detailed genetic and physical map for
RT   the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL   Mol. Microbiol. 21:77-96(1996).
DR   EMBL; AL589708; CAC33923.1; -.
DR   InterPro; IPR000461; Alpha_amylase.
DR   Pfam; PF00128; alpha-amylase; 1.
SQ   SEQUENCE   581 AA;  63952 MW;  8816DFF3BC1A3408 CRC64;


  Query Match             31.9%;  Score 942.5;  DB 2;  Length 581;
  Best Local Similarity   38.5%;  Pred. No. 7.5e-55;
  Matches  194;  Conservative   93;  Mismatches  174;  Indels   43;  Gaps   10;

Qy     15 FKLWAPYVNSVKLKLSKKLIPMEKNDE--GFFEVEIDDIEENLTYSYIIEDKREIPDPAS 72
Db      3 FEVWAPQAGRVTLRCDGATRALERDPERPGWWCGEA-PARDGSRYGFAVDDGPVLPDPRS 61

Qy     73 RYQPLGVHDKSQLIRTDYQILDLGKV---------KIEDLIIYELHVGTFSQEGNFKGVI 123
Db     62 RRQPDGPDGLS-------AVVDHGRYAWRTPWAGRPLPGGVLYELHVGTYTPEGTLDAAA 114

Qy    124 EKLDYLKDLGITGIELMPVAQFPGNRDWGYDGVFLYAVQNTYGGPWELAKLVNEAHKRGI 183
Db    115 DRLEHLVRLGVTHVELMPLCPFPGRHGWGYEGVSLWAVHEPYGGPEALKRFVDRAHDLGL 174

Qy    184 AVILDVVYNHIGPEGNYLLGLGPYFSDRYKTPWGLTFNFDDRGCDQVRKFILENVEYWFK 243
Db    175 GVVLDVVHNHLGPSGNHLPAFGPYFTDTHHTPWGVAVNLDAPGSDEVRAYLVDSALAWLR 234

Qy    244 TFKIDGLRLDAVHAIFDNSPKHILQEIAEKAHQLGK------FVIAESDLNDPKIV--KD 295
Db    235 DYRLDGLRLDAVHALADTRACHFLEQLSTAVDGLAADLERPLFLIAESDLNDPRIITPRA 294

Qy    296 DCGYKIDAQWVDDFHHAVHAFITKEKDYYYQDFGR--IEDIEKTFKDVFVYDGKYSRYRG 353
Db    295 EGGLGVHAQWNDDFHHALHAALTGESQGYYADFARDPLAALAKTLIRGYFHDGTYSSFRG 354

Qy    354 RTHGAPV--GDLPPRKFVVFIQNHDQVGNRGNGERLSILTDKTTYLMAATLYILSPYIPL 411
Db    355 RSHGRALDRGRVAAHRLTGYSQTHDQVGNRAQGDRLASLVSPGLAACAATLTLTAPFTPM 414

Qy    412 IFMGEEYYETNPFFFFSDFSDPVLIKGVREGRLKE-------NNQMIDPQSEEAFLKSKL 464
Db    415 LFMGEEWAAGTPWQFFTDHTDPQLAQAVRQGRRREFAAHGWREEDVPDPQDPATRERSCL 474

Qy    465 SWKIDE-----EVLDYYKQLINIR 483
Db    475 DWSEPEREPHARVLDWYRRLIALR 498


RESULT   7
Q9I1V1
ID   Q9I1V1      PRELIMINARY;      PRT;   583 AA.
AC   Q9I1V1;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE   PROBABLE GLYCOSYL HYDROLASE.
GN   PA2164.
OS   Pseudomonas aeruginosa.
OC   Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC   Pseudomonas.
OX   NCBI_TaxID=287;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 15692 / PAO1;
RX   MEDLINE=20437337; PubMed=10984043;
RA   Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrener P.,
RA   Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA   Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA   Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA   Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA   Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT   "Complete genome sequence of Pseudomonas aeruginosa PAO1, an
RT   opportunistic pathogen.";
RL   Nature 406:959-964(2000).
DR   EMBL; AE004643; AAG05552.1; -.
DR   InterPro; IPR002086; Aldehyde_dehydr.
DR   InterPro; IPR000461; Alpha_amylase.
DR   Pfam; PF00128; alpha-amylase; 1.
```

```
DR   PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
KW   Hydrolase; Complete proteome.
SQ   SEQUENCE   583 AA;  65657 MW;  651CD7A918833E00 CRC64;


  Query Match           29.3%;  Score 865.5;  DB 16;  Length 583;
  Best Local Similarity   38.2%;  Pred. No. 1e-49;
  Matches  199;  Conservative   89;  Mismatches  166;  Indels   67;  Gaps   20;

Qy       4 FGGNIEKN-KGIFKLWAPYVNSVKLKLSK-KLIPMEKNDEGFFEVEIDDIEENLTYSYII 61
           ||   : | :  | ||||    |::: :  :  |:| :| |:||  :           | | |
Db       7 FGAQFQGNGRTCFGLWAPDAREVRVETADGRDWPLEGSDSGWFEATL-PCPPGTRYRYRI 65

Qy      62 EDKREIPDPASRYQPLGVHDKSQLIRTDYQILDLGKV--KIE--------DLIIYELHVG 111
           : :  :||||||::|| ||| |       |:|| |    :::         : :|||||||
Db      66 DGRPGVPDPASQFQPDGVHGHS-------QVLDHGTYAWRVDEWRGRPWHEAVIYELHVG 118

Qy     112 TFSQEGNFKGVIEKLDYLKDLGITGIELMPVAQFPGNRDWGYDGVFLYAVQNTYGGPWEL 171
             |   |::  |   |  | :||:| :|||| :||| |:|||||  :|  : || | :|
Db     119 LF---GSYAEVERFLPRLVELGVTAVELMPLGEFPGRRNWGYDGVLPFAPASAYGTPEQL 175

Qy     172 AKLVNEAHKRGIAVILDVVYNHIGPEGNYLLG-LGPYFSDRYKTPWGLTFNFDDRGCDQV 230
             |:: ||  |: | :||:||| ||:|||||      :| |  :||||   :|  ||  :|
Db     176 KHLIDSAHGMGLMVFVDVIYNHFGPDGNYLAQYAAAFFRDDRQTPWGQAIDF-RRG--EV 232

Qy     231 RKFILENVEYWFKTFKIDGLRLDAVHAIFDNSPKHILQEIAEKAH-----QLGKFVIAES 285
           |:|  ||   |   :::|||| |||||| |::    | |:| :       :    :: |:
Db     233 REFFYENALMWLLDYRVDGLRFDAVHAIPDSA---FLVEMARRLRGAAGPERHVHLVLEN 289

Qy     286 DLNDPKIVKDDCGYKIDAQWVDDFHHAVHAFITKEKDYYYQDFGR-IEDIEKTFKDVFVY 344
           | |   :::   ||  |||| || |||:|  :| | | ||||:   :  : :   : |||
Db     290 DDNRASLLRQ--GY--DAQWNDDGHHALHVLLTGENDGYYQDYPEPLRCLARCLAEGFVY 345

Qy     345 DGKYSRYRGRTHGAPVGDLPPRKFVVFIQNHDQVGNRGNGERLSILTDKTTYLMAATLYI 404
           |: :|: ||  | |  || |  ||:|:|||||||||  |||||:| :     :|  | :
Db     346 QGEANRH-GRPRGEPSADLAPDAFVLFLQNHDQVGNRAFGERLSVLAEPQALRLAIALQL 404

Qy     405 LSPYIPLIFMGEEYYETNPFFFFSDFSDPVLIKGVREGRLKENNQ----------MID 452
           |:| |||:|||||     || :|:|     |   ||||| ||  :           : |
Db     405 LAPMIPLLFMGEECAAREPFLYFTDHQGE-LADAVREGRRKEFGEFGRFGEGATLASLPD 463

Qy     453 PQSEEAFLKSK-------LSWKIDEEVLDYYKQLINIRKRY 486
           | : | | :|:        :|:       :|:||: ||  :
Db     464 PNAVETFERSRPGLADCDPAWR------GFYRQLLEIRHEH 498


RESULT   8
Q9RX51
ID   Q9RX51      PRELIMINARY;      PRT;   600 AA.
AC   Q9RX51;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   MALTOOLIGOSYLTREHALOSE TREHALOHYDROLASE, PUTATIVE.
GN   DR0464.
OS   Deinococcus radiodurans.
OC   Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.
OX   NCBI_TaxID=1299;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=R1;
RX   MEDLINE=20036896; PubMed=10567266;
RA   White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA   Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA   Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA   Vamathevan J.J., Lam P., McDonald L., Utterback T., Zalewski C.,
RA   Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA   Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA   Fraser C.M.;
RT   "Genome sequence of the radioresistant bacterium Deinococcus
RT   radiodurans R1.";
RL   Science 286:1571-1577(1999).
DR   EMBL; AE001905; AAF10042.1; -.
DR   TIGR; DR0464; -.
DR   InterPro; IPR000461; Alpha_amylase.
DR   InterPro; IPR004193; isoamylase_N.
DR   Pfam; PF00128; alpha-amylase; 1.
DR   Pfam; PF02922; isoamylase_N; 1.
KW   Hydrolase; Complete proteome.
SQ   SEQUENCE   600 AA;  66909 MW;  594091EC093F8A44 CRC64;


  Query Match           28.3%;  Score 835.5;  DB 16;  Length 600;
  Best Local Similarity   36.8%;  Pred. No. 1.1e-47;
  Matches  181;  Conservative   89;  Mismatches  191;  Indels   31;  Gaps   12;

Qy      15 FKLWAPYVNSVKLKLSKKLIPMEKNDEGFFEVEIDDIEENLTYSYIIEDKREIPDPASRY 74
           |:||     :| ::::     |     | :|:|:  :     | ::: |    ||| :|:
Db      38 FRLWTSTARTVAVRVNGTEHVMTSLGGGIYELEL-PVGPGARYLFVL-DGVPTPDPYARF 95

Qy      75 QPLGVHDKSQLI---RTDYQILDLGKVKIEDLIIYELHVGTFSQEGNFKGVIEKLDYLKD 131
           | ||| :::::     |:   |   :|: | : ||:|||||: || ::   ||| |||:
Db      96 LPDGVHGEAEVVDFGTFDWTDADWHGIKLADCVFYEVHVGTFTPEGTYRAAAEKLPYLKE 155

Qy     132 LGITGIELMPVAQFPGNRDWGYDGVFLYAVQNTYGGPWELAKLVNEAHKRGIAVILDVVY 191
           ||:| |::||:| | | | |||||   ||    || | :|  ||: ||: |: | |||||
Db     156 LGVTAIQVMPLAAFDGQRGWGYDGAAFYAPYAPYGRPEDLMALVDAAHRLGLGVFLDVVY 215

Qy     192 NHIGPEGNYLLGLGP-YFSDRYKTPWGLTFNFDDRGCDQVRKFILENVEYWFKTFKIDGL 250
           || || ||||    | ||:||: : ||:  :: :     :|:::  |   | : :  |||
Db     216 NHFGPSGNYLSSYAPSYFTDRFSSAWGMGLDYAE---PHMRRYVTGNARMWLRDYHFDGL 272

Qy     251 RLDAVHAIFDNSPKHILQEIAEKAHQLG--KFVIAESDLNDPKIVKDDCGYKIDAQWVDD 308
           ||||   : |:|  ||| |:|:: |:||    ::||   | | :|  :   :|  | ||
Db     273 RLDATPYMTDDSETHILTELAQEIHELGGTHLLLAEDHRNLPDLVTVN---HLDGIWTDD 329

Qy     309 FHHAVHAFITKEKDYYYQDF-GRIEDIEKTFKDVFVYDGKYSRYRGRTH--GAPVGDLPP 365
           |||     :| |:: ||  : |  | :  | :  : |:|::   :|  |  | |   |
Db     330 FHHETRVTLTGEQEGYYAGYRGGAEALAYTIRRGWRYEGQFWAVKGEEHERGHPSDALEA 389

Qy     366 RKFVVFIQNHDQVGNRGNGERLSILTDKTTYLMAATLYILSPYIPLIFMGEEYYETNPFF 425
             ||  ||||||:|||  ||||      | :       :| |  ||:| |:|:  : ||
Db     390 PNFVYCIQNHDQIGNRPLGERLHQSDGVTLHEYRGAAALLLPMTPLLFQGQEWAASTPFQ 449

Qy     426 FFSDFSDPVLIKGVREGRLKE--------NNQMIDPQSEEAFLKSKLSWKIDE-----EV 472
           |||| :   | : | ||| ||           : |||:|: || |||:)   |
Db     450 FFSDHAGE-LGQAVSEGRKKEFGGFSGFSGEDVPDPQAEQTFLNSKLNWAEREGGEHART 508

Qy     473 LDYYKQLINIRK 484
           |  |: |: :|:
Db     509 LRLYRDLLRLRR 520


RESULT   9
Q92U63
ID   Q92U63      PRELIMINARY;      PRT;   601 AA.
AC   Q92U63;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PUTATIVE 1,4-ALPHA-GLUCAN BRANCHING ENZYME PROTEIN (EC
DE   2.4.1.18).
OS   Rhizobium meliloti (Sinorhizobium meliloti).
OG   Plasmid pSymB (megaplasmid 2).
OC   Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC   Rhizobiaceae; Sinorhizobium.
OX   NCBI_TaxID=382;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=1021;
RX   MEDLINE=21396508; PubMed=11481431;
RA   Finan T.M., Weidner S., Wong K., Buhrmester J., Chain P.,
RA   Vorhoelter F.J., Hernandez-Lucas I., Becker A., Cowie A., Gouzy J.,
RA   Golding B., Puehler A.;
RT   "The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-
RT   fixing endosymbiont Sinorhizobium meliloti.";
```

```
RL   Proc. Natl. Acad. Sci. U.S.A. 98:9889-9894(2001).
DR   EMBL; AL603646; CAC49676.1; -.
KW   Transferase; Glycosyltransferase; Plasmid; Hypothetical protein;
KW   Complete proteome.
SQ   SEQUENCE   601 AA;  66954 MW;  024DA322A7B72C2E CRC64;


  Query Match             26.8%;  Score 792.5;  DB 16;  Length 601;
  Best Local Similarity   36.9%;  Pred. No. 8e-45;
  Matches  192;  Conservative   79;  Mismatches  199;  Indels   51;  Gaps   15;

Qy       3 SFGGN-IEKNKGIFKLWAPYVNSVKLKLSKKLIPMEKNDEGFFEVEIDDIEENLTYSYII 61
           |:| | |: :   |:||||    | | |   :  |:  | |:||: :   :    | : :
Db      13 SWGANFIDSDTCRFRLWAPDEREVDLVLGGAVHKMQSLDGGWFEITL-AAKAGERYCFRL 71

Qy      62 EDKREIPDPASRYQPLGVHDKSQLI---RTDYQILDLGKVKIEDLIIYELHVGTFSQEGN 118
             |  |: ||||  |       | ::    ::|       |: :| ||||| |: ||
Db      72 ADGTEVADPASSAQEREASGTSIVVDHAAYEWQTSSWRGRPWEEAVISELHVGCFTPEGT 131

Qy     119 FKGVIEKLDYLKDLGITGIELMPVAQFPGNRDWGYDGVFLYAVQNTYGGPWELAKLVNEA 178
           |:  ||:| :|   ||| ||:|||||||| | ||||||  ||  | || | :|  ||: |
Db     132 FRAAIERLPHLAGAGITAIEIMPVAQFPGVRGWGYDGVLHYAPHNAYGKPDDLKALVDAA 191

Qy     179 HKRGIAVILDVVYNHIGPEGNYLLGLGP-YFSDRYKTPWGLTFNFDDRGCDQVRKFILEN 237
           |  |: |:|||||| |||||||      :|:    |||| :  |::   : ||:: :||
Db     192 HSLGLTVLLDVVYNHFGPEGNYLSRYASRFFNKDRPTPWGASIAFEE---EAVRRYFIEN 248

Qy     238 VEYWFKTFKIDGLRLDAVHAIFDNSPKHIL----QEIAEK-AHQLGKFVIAESDLNDPKI 292
             ||   |: |||||||   | | :  | |     |: |  | :    |: ::      :
Db     249 ALYWLGHFRFDGLRLDATEQIRDTTKPHFLVALEHEVREAFAERQIHLVLEDAHRRRSLL 308

Qy     293 VKDDCGYKI--DAQWVDDFHHAVHAFITKEKDYYYQDF-----GRIEDIEKTFKDVFVYD 345
             :|  | ::  || | || |:|:|   | |   :|: |     |:   |     : |
Db     309 QRDASGARMLFDAAWNDDLHNALHVVATGETKGHYRLFADEPWGK---IRSALAEGFAVP 365

Qy     346 GKYSRYRGRTHGAPVGD---LPPRKFVVFIQNHDQVGNRGNGERLSILTDKTTYLMAATL 402
             |   :      :| |    :||:  | |:|||||:|||  ||||: |  : :  : | :
Db     366 AKEDNF------SPEGSRARVPPQGRVNFLQNHDQIGNRAFGERLASLLQEDSLRVLAAM 419

Qy     403 YILSPYIPLIFMGEEYYETNPFFFFSDFSDPVLIKGVREGRL-----------KENNQM 450
           ::|:| |||:|||||| || ||:||||:   :   :| ||              |  : :
Db     420 HMLTPQIPLLFMGEEYGETQPFYFFSDYQGEI-AAAIRLGRRDEAENFGGLPEGKTMDDL 478

Qy     451 IDPQSEEAFLKSKLSWK-----IDEEVLDYYKQLINIRKRY 486
             ||   : |  ||| |       |  | | : |  ||:|:
Db     479 PDPLDPDVFAGSKLRWNRATSPAGERHLAYVRDLAVIRQRH 519


RESULT  10
Q973H5
ID   Q973H5      PRELIMINARY;      PRT;   217 AA.
AC   Q973H5;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PUTATIVE ALPHA-AMYLASE.
GN   ST0926.
OS   Sulfolobus tokodaii.
OC   Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.
OX   NCBI_TaxID=111955;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JCM 10545 / 7;
RX   PubMed=11572479;
RA   Kawarabayasi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
RA   Sekine M., Baba S.-I., Ankai A., Kosugi H., Hosoyama A., Fukui S.,
RA   Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
RA   Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,
RA   Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,
RA   Oshima T., Kikuchi H.;
RT   "Complete genome sequence of an aerobic thermoacidophilic
RT   Crenarchaeon, Sulfolobus tokodaii strain7.";
RL   DNA Res. 8:123-140(2001).
DR   EMBL; AP000984; BAB65938.1; -.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   217 AA;  25181 MW;  4A1235EC5B491CEF CRC64;


  Query Match             21.8%;  Score 645;  DB 17;  Length 217;
  Best Local Similarity   58.9%;  Pred. No. 1.4e-35;
  Matches  122;  Conservative   37;  Mismatches   44;  Indels    4;  Gaps    4;

Qy       4 FGGN-IEKNKGIFKLWAPYVNSVKLKLSKKLIPMEKNDEGFFEVEIDDIEENLTYSYIIE 62
           || : : :|:  | |||||     |:|:::   |||:::|:| : : : :|   ||:||
Db       2 FGPSFLSENEVEFILWAPYQKKATLRLNEEKYDMEKDEKGYFRITV-NAKEGDKYSFII- 59

Qy      63 DKREIPDPASRYQPLGVHDKSQLIRTDYQILDLG-KVKIEDLIIYELHVGTFSQEGNFKG 121
           || |||||||||| ||| :|||   |:      |:| |||||||||:|||:||| |
Db      60 DKGEIPDPASRYQPDGVHGRSQLISLKYEWKSREVKIKPEDLIIYELHIGTFTQEGTFYS 119

Qy     122 VIEKLDYLKDLGITGIELMPVAQFPGNRDWGYDGVFLYAVQNTYGGPWELAKLVNEAHKR 181
            :|||||| :||:| ||:||| |||| :|||||||:|||||||||||:   | ::|||:|
Db     120 AVEKLDYLVNLGVTAIEIMPVHQFPGKKDWGYDGVYLYAVQNTYGGPYAFMKFIDEAHQR 179

Qy     182 GIAVILDVVYNHIGPEGNYLLGLGPYF 208
           |:|||||||||:|||||||:: ||||:
Db     180 GLAVILDVVYNHVGPEGNYMMFLGPYY 206


RESULT  11
Q44528
ID   Q44528      PRELIMINARY;      PRT;   552 AA.
AC   Q44528;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PUTATIVE  ALPHA-GLUCANOTRANSFERASE /HYDROLASE.
GN   LTI2.
OS   Anabaena variabilis.
OC   Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Anabaena.
OX   NCBI_TaxID=1172;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=M3;
RX   MEDLINE=92119230; PubMed=1731972;
RA   Sato N.;
RT   "Cloning of a low-temperature-induced gene lti2 from the
RT   cyanobacterium Anabaena variabilis M3 that is homologous to alpha-
RT   amylases.";
RL   Plant Mol. Biol. 18:165-170(1992).
DR   EMBL; X59279; CAA41970.1; -.
DR   InterPro; IPR000461; Alpha_amylase.
DR   InterPro; IPR004193; isoamylase_N.
DR   Pfam; PF00128; alpha-amylase; 1.
DR   Pfam; PF02922; isoamylase_N; 1.
KW   Transferase; Hydrolase.
SQ   SEQUENCE   552 AA;  63833 MW;  069886FAFD8FC411 CRC64;


  Query Match             16.7%;  Score 493;  DB 2;  Length 552;
  Best Local Similarity   28.6%;  Pred. No. 6.8e-25;
  Matches  159;  Conservative   76;  Mismatches  142;  Indels  178;  Gaps   26;

Qy      15 FKLWAPYVNSVKLKLS---KKLIPMEKNDEGFFE--VEIDDIEENL-----TYSYIIEDK 64
             | |:|||     |  |    : |||:| |:|:|   ||::|          | |:  |:
Db       7 FTLFAPYNKGAALIASFSDWQEIPMKKGDDGYFRTTVELEDGTYQYKFRVQTRSWFFEED 66

Qy      65 R--EIPDPASR--YQPLGVHDKSQLIRTDYQILDLGKVKIED--------LIIYELHVGT 112
           :  :: || :    :  |  :    |:   :|:|    : :|        |:|||||||
Db      67 QWVDVTDPYATDIDESSGKDNSIARIKDGEKIVDTYVWQHDDKPLPADHELVIYELHVGD 126

Qy     113 FS-------QEGNFKGVIEKLDYLKDLGITGIELMPVAQFPGNRDWGYDGVFLYAVQNTY 165
           ||         | :| |||||||| :|||  |||:|| ::||:  |||:  : :| :::|
Db     127 FSGGEDDPYARGKYKHVIEKLDYLCELGINAIELLPVKEYPGDYSWGYNPRYFFATESSY 186

Qy     166 GGPWELAKLVNEAHKRGIAVILDVVYNHI-------------------GPEGNYLLGLG 205
```

```
          |   :| |||:| |:||| :|:| :|||                          |: |
Db    187 GSTADLKKLVDECHQRGIRIIMDGIYNHSEASSPLTQIDHDYWYHHEPRDPDNN------ 240

Qy    206 PYFSDRYKTPWGLTFNFD--DRGCD--QVRKFILENVEYWFKTFKIDGLRLDAVHAIFDN 261
                ||  ||:: |   :    |||| : | ||   : :||:| ||   | :
Db    241 ----------WGPEFNYEHYDENLETYPARKFIGDTVRYWVGEYHLDGIRYDAARQIANY 290

Qy    262 SPKH-ILQEIAEKAHQLGKFVIAE--------SDLNDPKIVKDDCGYKIDAQWVDDFHHA 312
          | | ||  : |     : :||        ::|: |          :|  | | |:|
Db    291 DFMHWIAQEAKKTAGAKPFYNVAEHIPETTSITNLDGP----------MDGCWHDSFYHT 340

Qy    313 VHAFITKEKDYYYQDFGRIEDIEKTFKDVFVYDGKYSRYRGRTHGAPVGDLPPRKFVVFI 372
          : | |           |   |:|   |||   | |   : | |:          | ::
Db    341 IKAHIC----------GDTFDLE-NLKDVI--DPKRQGFLGATN----------VVNYL 376

Qy    373 QNHD------QVGNRGNGERLSILTDK--------TTYLMAATLYILSPYIPLIFMGEEY 418
           |||      ::|||       |  |:         |  || |        :|||:||||:
Db    377 TNHDHDHIMVELGNR------EIFHDEAFRRAKLGTAILMTAV------GVPLIWMGEEF 424

Qy    419 YETNPFFFFSDFSDPVLIKGVREGRLKENNQMIDPQSEEAFLKSKLSW-----KIDEEVL 473
          |  |                      |: :|            ||: |      ::  :
Db    425 GEYKP--------------------KQQDQ-----------SKIDWTLLGNDLNRSLF 451

Qy    474 DYYKQLINIRKRYNN 488
          ||:| || :||  ||
Db    452 DYHKGLIGLRK--NN 464


RESULT  12
Q93Q35
ID   Q93Q35      PRELIMINARY;      PRT;   440 AA.
AC   Q93Q35;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   BRANCHING ENZYME GLGB (FRAGMENT).
OS   Myxococcus xanthus.
OC   Bacteria; Proteobacteria; delta subdivision; Myxobacteria;
OC   Myxococcales; Cystobacterineae; Myxococcaceae; Myxococcus.
OX   NCBI_TaxID=34;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Ueki T., Inouye S.;
RT   "Identification of a new His-Asp phosphorelay signal transduction
RT   system which regulates expression of a heat shock gene, lonD, of
RT   Myxococcus xanthus.";
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF285783; AAK83002.1; -.
FT   NON_TER     440    440
SQ   SEQUENCE   440 AA;  48964 MW;  8D492EA5A4A92017 CRC64;


  Query Match             14.2%;  Score 420;  DB 2;  Length 440;
  Best Local Similarity   32.0%;  Pred. No. 3.7e-20;
  Matches  133;  Conservative   52;  Mismatches  129;  Indels  102;  Gaps   17;

Qy     52 EENLTYSYIIEDKREIPDPASRYQPLGVHDKSQLIRTDYQILDLGKVKIEDLIIYELHVG 111
          :|| |'| || |  |    | :|  |                ::||||:|||
Db     80 QENSTGSSIIYDHGEYWWNAQQYSSPG--------------------FNEMIIYEMHVG 118

Qy    112 TFSQE-----GNFKGVIEKLDYLKDLGITGIELMPVAQFPGNRDWGYDGVFLYAVQNTYG 166
          ||        ||:   | :||:::|||   |::||  :| |:  |||:  | :| :: ||
Db    119 TFHDSPGFGPGNWNSAIARLDHVRDLGANMIKVMPAYEFAGDFSWGYNAAFPFAPESAYG 178

Qy    167 GPWELAKLVNEAHKRGIAVILDVVYNHIGP--------EGNYLLGLGPYFSD--RYKTPW 216
          | :: : |:||| ||| || |||:|| ||          |: |   | || :  |  |||
Db    179 HPNDMKRFVDEAHMRGIGVIFDVVHNHYGPSDLPMWCFSGDCLGSGGEYFYNDWRKSTPW 238

Qy    217 GLTFNFDDRGCDQVRKFILENVEYWFKTFKIDGLRLDAVH----------AIFD--NSP 263
          | |    | |  :|| :| :::     :|: |||| ||            || |
Db    239 GDT--RPDYGRPEVRAYIRDSMMNLTHSFRGDGLRWDATKYMRTQNGSDTTAIPDAWRVF 296

Qy    264 KHILQEIAEKAHQLGKFVIAESDLNDPKIVKDDC------GYKIDAQWVDDFHHAVHAFI 317
          : | :||   | |  |  ||| |      : :|       |   |:||  || ||: | :
Db    297 RSINREI--NATQPWKISIAE-DFGGGDFITNDATSDTSGGAGFDSQWGGDFVHAIRAAV 353

Qy    318 TKEKDYYYQDFGRIEDIEKTFKDVFVYDGKYSRYRGRTHGAPVGDLPPRKFVVFIQNHDQ 377
                  | ||  |:        | :    || || | |          |:: ::||:
Db    354 IAS-----NDSGR--DMNS------VRNAITQRYSGR-HTAR---------VIYSESHDE 390

Qy    378 V-------------GNRGNGERLSILTDKTTYLMAATLYILSPYIPLIFMGEEYYE 420
          |              || |:         |    :|| :   || ||:|| |:|: |
Db    391 VANGKARVPEEIWPGNAGS------WAAKKRSTLAAGVVFTSPGIPMIFQGQEFLE 440


RESULT  13
Q59242
ID   Q59242      PRELIMINARY;      PRT;   652 AA.
AC   Q59242; O08511; O08485;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   1,4-ALPHA-GLUCAN BRANCHING ENZYME (EC 2.4.1.18)
DE   (GLYCOGEN BRANCHING ENZYME) (1,4-ALPHA-GLUCAN BRANCHING ENZYME)
DE   (AMYLO-(1,4 TO 1,6)TRANSGLUCOSIDASE)
DE   (AMYLO-(1,4-1,6)-TRANSGLYCOSYLASE).
GN   GLGB.
OS   Bacillus stearothermophilus.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Bacillus/Staphylococcus group; Geobacillus.
OX   NCBI_TaxID=1422;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TRBE14;
RX   MEDLINE=95031021; PubMed=7944355;
RA   Takata H., Takaha T., Kuriki T., Okada S., Takagi M., Imanaka T.;
RT   "Properties and active center of the thermostable branching enzyme
RT   from Bacillus stearothermophilus.";
RL   Appl. Environ. Microbiol. 60:3096-3104(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TRBE14;
RA   Takata H.;
RL   Submitted (AUG-1996) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TRBE14;
RA   Takata H., Takaha T., Okada S., Takagi M., Imanaka T.;
RL   Submitted (APR-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- CATALYTIC ACTIVITY: FORMATION OF 1,6-GLUCOSIDIC LINKAGES OF
CC       GLYCOGEN.
CC   -!- PATHWAY: THIRD STEP IN GLYCOGEN BIOSYNTHESIS.
CC   -!- MISCELLANEOUS: OPTIMAL ACTIVITY AT APPROXIMATELY 39 DEGREES
CC       CELSIUS.
CC   -!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC       KNOWN AS THE ALPHA-AMYLASE FAMILY.
DR   EMBL; D87026; BAA19588.1; -.
DR   InterPro; IPR000461; Alpha_amylase.
DR   InterPro; IPR004193; isoamylase_N.
DR   Pfam; PF00128; alpha-amylase; 1.
DR   Pfam; PF02922; isoamylase_N; 1.
KW   Glycogen biosynthesis; Transferase; Glycosyltransferase.
FT   ACT_SITE    308    308       BY SIMILARITY.
FT   ACT_SITE    351    351       BY SIMILARITY.
FT   ACT_SITE    419    419       BY SIMILARITY.
SQ   SEQUENCE   652 AA;  76793 MW;  4591EB414A0E3FEF CRC64;


  Query Match             12.2%;  Score 360;  DB 2;  Length 652;
  Best Local Similarity   25.3%;  Pred. No. 6.3e-16;
  Matches  142;  Conservative   72;  Mismatches  186;  Indels  162;  Gaps   25;

Qy      4 FGGNIEKNKG----IFKLWAPYVNSVKL-----------KLSKKLIPMEKNDEGFFEVE 47
          || ::    |     | :|||:   |:|           :| |       |||| : :
Db     26 FGAHVINEGGKVGTRFCVWAPHAREVRLVGSFNDWDGTDFRLEK------VNDEGVWTIV 79
```

```
Qy      48 IDDIEENLTYSY-IIEDKREI---PDPASRYQPLGVHDKSQLIRTD---YQILDLG---- 96
Db      80 VPENLEGHLYKYEIVTPDGQVLFKADPYAFYSELRPHTAS--IAYDLKGYQWNDQSWKRK 137

Qy      97 ----KVKIEDLIIYELHVGTF-SQEGNFKGVIEKLD----YLKDLGITGIELMPVAQFPG 147
Db     138 KRRKRIYDQPMVIYELHFGSWKKKDGRFYTYREMADELISYVLDHGFTHIELLPLVEHPL 197

Qy     148 NRDWGYDGVFLYAVQNTYGGPWELAKLVNEAHKRGIAVILDVVYNHI--GPEGNYLLGLG 205
Db     198 DRSWGYQGTGYYAVTSRYGTPHDFMYFVDRCHQAGIGVIMDWVPGHFCKDAHGLYMFDGA 257

Qy     206 PYF-----SDRYKTPWGLTFNFDDRGCDQVRKFILENVEYWFKTFKIDGLRLDAV----- 255
Db     258 PTYEYANEKDRENYVWG-TANF-DLGKPEVRSFLISNALFWLEYYHIDGFRVDAVANMLY 315

Qy     256 ----HAIFDNS-PKHILQEIAEK--AHQLGKFVIAESDLNDPKIVKD------DCGYKID 302
Db     316 WPNNDRLYENPYAVEFLRKLNEAVFAYDPNALMIAEDSTDWPKVTAPTYEGGLGFNYKWN 375

Qy     303 AQWVDDF----------HHAVHAFITKEKDYYYQDFGRIEDIEKTFKDVFVYDGKYSRYR 352
Db     376 MGWMNDMLKYMETPPYERRHVHNQVTFSLLYAYS-----ENFILPFSHDEVVHGKKS--- 427

Qy     353 GRTHGAPVGDLPPRKFVVFIQNHDQVGNRGNG---ERLSILTDKTTYLMAATLYILSPYI 409
Db     428 ------------------------LLNKMPGSYEEKFAQLRLLYGYMMA------HPGK 456

Qy     410 PLIFMGEEYYETNPFFFFSDFSDPVLIKGVREGRLKENNQMIDPQSEEAFLKSKLSW--- 466
Db     457 KLLFMGNEF----------------------------------AQFDEWKFEDELDWVLF 482

Qy     467 --KIDEEVLDYYKQLINIRKRY 486
Db     483 DFELHRKMNDYMKELIACYKRY 504
```

```
RESULT  14
P73608
ID   P73608      PRELIMINARY;      PRT;   707 AA.
AC   P73608;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   GLYCOGEN OPERON PROTEIN GLGX.
GN   GLGX OR SLR1857.
OS   Synechocystis sp. (strain PCC 6803).
OC   Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX   NCBI_TaxID=1148;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97061201; PubMed=8905231;
RA   Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA   Miyajima N., Hirosawa M., Sugiura M., Sasamoto S., Kimura T.,
RA   Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
RA   Shimpo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
RA   Tabata S.;
RT   "Sequence analysis of the genome of the unicellular cyanobacterium
RT   Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT   entire genome and assignment of potential protein-coding regions.";
RL   DNA Res. 3:109-136(1996).
DR   EMBL; D90908; BAA17652.1; -.
DR   HSSP; P10342; 1BF2.
DR   InterPro; IPR000461; Alpha_amylase.
DR   InterPro; IPR004193; isoamylase_N.
DR   Pfam; PF00128; alpha-amylase; 1.
DR   Pfam; PF02922; isoamylase_N; 1.
KW   Complete proteome.
SQ   SEQUENCE   707 AA;  79895 MW;  CFA27B3C86318DB1 CRC64;
```

```
  Query Match           11.3%;  Score 333.5;  DB 16;  Length 707;
  Best Local Similarity  23.1%;  Pred. No. 4.1e-14;
  Matches  146;  Conservative   83;  Mismatches  197;  Indels  207;  Gaps   29;

Qy       2 FSFGGNIEKNKGIFKLWAPYVNSVKLKLSKKLIPMEKNDEGFFEVE---------- 47
Db      23 FPFGATIVPGGVNFSIYSSHSTACTLVLFEKRAP-----QPFVEIPFPESFRIGNVYCMV 77

Qy      48 IDDIE-ENLTYSYIIEDKREIPDPASRYQPLGVHDKSQLIRTDYQILDLGK--------- 97
Db      78 VFDLDFENLEYGYRMEG-------PNNFQQGHWFDPSKVLLDPYAKVVSGRDVWGTQPNW 130

Qy      98 ------------------------VKIEDLIIYELHVGTFSQE----------GNFKGVI 123
Db     131 DDIYQHRGRLSFDDFDWENDSPLDVPLEDMVIYEMHVRGFTKDPSSGVKENHRGTFAGIL 190

Qy     124 EKLDYLKDLGITGIELMPVAQF---------PGNRD-----WGYDGVFLYAVQNTYGGPW 169
Db     191 SKIPYLQELGVNTIELMPIFEFDEFEHSRYHPETGEFLVNYWGYSTVNFFAPKAGYAATG 250

Qy     170 -------ELAKLVNEAHKRGIAVILDVVYNHIGPEGNYLLGLGPYFSDR--------YKT 214
Db     251 KFGMQIDELKNLVKELHKVGISVILDVVFNHTA-EGN---ERGPTISFRGLDNKTYYMLT 306

Qy     215 PWGLTFNFDDRG----CDQ--VRKFILENVEYWFKTFKIDGLRLDAVHAIFDNSPKHILQ 268
Db     307 PEGYYFNFSGTGNTLNCNNPIVRGMVLDCLRYWTAEFHIDGFRFD-LASILGRDP----- 360

Qy     269 EIAEKAHQLGKFVIAESDLNDP-----KIVKD--DCG--YKID--------AQWVDDFHH 311
Db     361 ----WGYPLANPPLLETLAFDPILARSKLIAEAWDAGGLYQVGSFPSYGRWAEWNGKYRD 416

Qy     312 AVHAFITKEKDY-------------YYQDFGRIEDIEKTF---------KDVFVYDGKYS 349
Db     417 TVRKFIKGDAGVIGEMAQRLQGSPDLYQGAGRPPSTSINFVTAHDGFTLADLVAYNGKHN 476

Qy     350 RYRGRTHGAPVGDLPPRKFVVFIQNHD-QVGNRGNGERLSILTDKTTYLM-AATLYILSP 407
Db     477 YANGENGNDGAND-----------NYSWNCGVEGPTDNPDILRLRARQMRNAIAILLVSQ 525

Qy     408 YIPLIFMGEE-----------YYETNPFFFFSDFSDPVLIKGVREGRLKENNQMIDPQSE 456
Db     526 GVPMLLMGDEMGKTQDGNNNTYCHDSPFNWLN------------WHLLEQN-----QAW 567

Qy     457 EAFLKSKLSWKIDEEVLDYYKQLINIRKRYNNC 489
Db     568 FRFVKHCIAFRLAHPVL-------RNSEHFQNC 593
```

```
RESULT  15
Q97SQ7
ID   Q97SQ7      PRELIMINARY;      PRT;  1280 AA.
AC   Q97SQ7;
DT   01-OCT-2001 (TrEMBLrel. 18, Created)
DT   01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   ALKALINE AMYLOPULLULANASE, PUTATIVE.
GN   SP0268.
OS   Streptococcus pneumoniae.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC   Streptococcus.
OX   NCBI_TaxID=1313;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TIGR4;
RX   MEDLINE=21357209; PubMed=11463916;
RA   Tettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
RA   Peterson S., Heidelberg J., DeBoy R.T., Haft D.H., Dodson R.J.,
RA   Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,
RA   Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
RA   Holtzapple E., Khouri H., Wolf A.M., Utterback T.R., Hansen C.L.,
RA   McDonald L.A., Feldblyum T.V., Angiuoli S., Dickinson T., Hickey E.K.,
RA   Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
RA   Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
RT   "Complete genome sequence of a virulent isolate of Streptococcus
```

```
RT   pneumoniae.";
RL   Science 293:498-506(2001).
DR   EMBL; AE007340; AAK74446.1; -.
DR   TIGR; SP0268; -.
DR   InterPro; IPR000461; Alpha_amylase.
DR   InterPro; IPR001899; Gram_pos_anchor.
DR   InterPro; IPR004193; isoamylase_N.
DR   Pfam; PF00128; alpha-amylase; 1.
DR   Pfam; PF00746; Gram_pos_anchor; 1.
DR   Pfam; PF02922; isoamylase_N; 1.
KW   Complete proteome.
SQ   SEQUENCE   1280 AA;  142619 MW;  A209B802A8E0B328 CRC64;


  Query Match           11.3%;  Score 332.5;  DB 16;  Length 1280;
  Best Local Similarity  21.9%;  Pred. No. 1.1e-13;
  Matches  152;  Conservative   96;  Mismatches  208;  Indels  237;  Gaps   32;

Qy       2 FSFGGNI-----EKNKGI-FKLWAPYVNSVKLKLSKK--------LIPMEKNDEGFFEVE 47
           :|: |.:     |: | :   ||:|  : | : :  |         : :|| : | ::
Db     445 YSYDGKLGADLKEEGKQVDLTLWSPSADKVSVVVYDKNDPDKVVGTVALEKGERGTWKQT 504

Qy      48 IDDIEE-------NLTYSYIIEDKREIP---DPASRYQPLGVHDKSQL------------ 85
           :|   :          | | || : :     || ::       | |::
Db     505 LDSTNKLGITDFTGYYYQYQIERQGKTVLALDPYAKSLAAWNSDDSKIDDAHKVAKAAFV 564

Qy      86 --IRTDYQILDLGKV----KIEDLIIYELHVGTFSQE-----------GNFKGVIEKLDY 128
             :   | |  ||:       || :||| ||  |: :           | |:  ||||||
Db     565 DPAKLGPQDLTYGKIHNFKTREDAVIYEAHVRDFTSDPAIAKDLTKPFGTFEAFIEKLDY 624

Qy     129 LKDLGITGIELMPV-----------------AQFPGNRDWGYDGVFLYAVQNTYGGP-- 168
           |||||:| |:|:||                 |    | :|||    :::   |
Db     625 LKDLGVTHIQLLPVLSYYFVNELKNHERLSDYASSNSNYNWGYDPQNYFSLTGMYSSDPK 684

Qy     169 ------WELAKLVNEAHKRGIAVILDVVYNHIGP-------EGNYLLGLGPYFSDRYKTP 215
                |   |:|| ||||:  ||||||||          | ||      :| |   ||
Db     685 NPEKRIAEFKNLINEIHKRGMGAILDVVYNHTAKVDLFEDLEPNYY-----HFMDADGTP 739

Qy     216 --------WGLTFNFDDRGCDQVRKFILENVEYWFKTFKIDGLRLDAVHAIFDNSPKHIL 267
                 | | :        :: ::::::|   |:|:|| |       ||    |
Db     740 RTSFGGGRLGTTHH-------MTKRLLIDSIKYLVDTYKVDGFR-------FDMMGDHDA 785

Qy     268 QEIAEKAHQLGKFVIAESDLNDPKIVKDDCGYK-----------IDAQWV--------- 306
             | |:|::  : :       :| ::    |::           |  |:
Db     786 ASI-EEAYKAARAL-------NPNLIMLGEGWRTYAGDENMPTKAADQDWMKHTDTVAVF 837

Qy     307 -DDFHHAVH---------AFITKEKDYYYQDFGRIEDIEKTFKDVFVY--------DGKY 348
            ||  : :          ||||          |   |:   ||::            |
Db     838 SDDIRNNLKSGYPNEGQPAFIT----------GGKRDVNTIFKNLIAQPTNFEADSPGDV 887

Qy     349 SRYRGRTHGAPVGDL-------PPRKFVVFIQNHDQVGNRGNGERLSILTDKTTYLMAAT 401
            :|       : |:       | |     :|: ::  |      | :|| : |
Db     888 IQYIAAHDNLTLFDIIAQSIKKDPSK----AENYAEIHRRLRLGNLMVLTAQGT------ 937

Qy     402 LYILSPYIPLIFMGEEYYETNPFFFFSDFSDPVLIKGVREGRLKENNQMIDPQSEEAF-- 459
                | |  |:||  |        | ||     | | ::   : ::  :    |
Db     938 --------PFIHSGQEYGRT------KQFRDPAYKTPVAEDKVPNKSHLLRDKDGNPFDY 983

Qy     460 ------------LKSKLSW-------KIDEEV--LDYYKQLINIRKRYN--NCKRVKEVR 496
                       :|  |          | |   || | || :|:  :    | :::::
Db     984 PYFIHDSYDSSDAVNKFDWTKATDGKAYPENVKSRDYMKGLIALRQSTDAFRLKSLQDIK 1043

Qy     497 REGNCITLI----MEKIGIIASFDDIVINSKIT 525
             : ||:     :||        :|:||  :||
Db    1044 DRVHLITVPGQNGVEK-------EDVVIGYQIT 1069


Search completed: July 15, 2002, 12:24:10
Job time: 392 sec
```